2014

The medicinal chemistry development for new antimicrobial chemotherapeutics

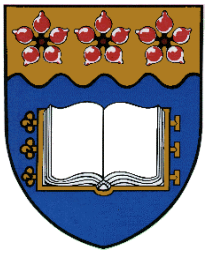
Adel Ahmed Rashad Ahmed
University of Wollongong

THE MEDICINAL CHEMISTRY DEVELOPMENT FOR NEW ANTIMICROBIAL CHEMOTHERAPEUTICS

A thesis submitted in fulfilment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY

From



UNIVERSITY OF WOLLONGONG

By

Adel Ahmed Rashad Ahmed

B. Pharm. Sci.(2005),

MSc. Med. Chem.(2009)

Supervisor: Assoc. Prof. Paul Keller

SCHOOL OF CHEMISTRY

August 2014

Certification

I, Adel Ahmed Rashad, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Chemistry, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

Adel Ahmed Rashad Ahmed

August 2014

Table of Contents

Certification i
Table of Contents ii
Publications v
Acknowledgements vi
Abstract vii
List of Abbreviations x
CHAPTER 1: Introduction 1
1.1. Classification, history and clinical features 1
1.2. Virology of the CHIKV 3
1.3. The development of CHIKV vaccine 6
1.4. Emerging novel CHIKV targets 7
1.4.1. Non-structural proteins 8
1.4.1.1. Non-structural protein 1 8
1.4.1.2. Non-structural protein 2 8
1.4.1.3. Non-structural protein 3 12
1.4.1.4. Non-structural protein 4 14
1.4.2. Structural proteins 15
1.5. Highlights for the CHIKV emerging targets 19
1.6. Development of chemotherapeutics against CHIKV: new medicinal chemistry leads 20
1.6.1. Protease inhibitors 20
1.6.2. Furin inhibitors 22
1.6.3. Chloroquine and Quinine 23
1.6.4. Ribavirin and 6-Azauridine 24
1.6.5. Arbidol 25
1.6.6. Mycophenolic acid (MPA) 26
1.6.7. Trigocherrin A 27
1.6.8. Trigowiin A, Prostratin and 12-O-Tetradecanoylphorbol 13-Acetate 27
1.6.9. Lupenone and β-amyrone 28
1.6.10. Harringtonine 29
1.6.11. Purine based inhibitors 29
1.6.12. Polyinosinic acid 30
1.6.13. Gene silencers 31

1.7. Highlights for CHIKV inhibitors 31
1.8. Concluding remarks 33
1.9. Project aims 34
CHAPTER 2: The Search for Anti-CHIKV Lead Compounds 35
2.1. Introduction 35
2.2. Synthesis of the *meta*-nitro AAPM derivative 37
2.3. Modification in the amine side chain of the AAPM derivatives 46
2.4. Decreasing the steric hindrance of the dimethyl groups 48
2.5. Replacing the 4,6-dichloropyrimidine with 4-chloropyridine 50
2.6. Concluding remarks 62
CHAPTER 3: Modelling Study of the CHIKV nsP3 as a Possible Drug Development Target 64
3.1. Introduction 64
3.2. Results and discussion: 66
3.3. Synthesis of pose_3_2 (**86**) 77
3.4. Concluding remarks 80
CHAPTER 4: Synthesis of Mycophenolic Acid Analogues as Inhibitors of the Chikungunya Virus 83
4.1. Introduction 83
4.2. Design of MPA analogues 84
4.3. Synthesis of the aromatic heads 86
4.3.1. Synthesis of the isobenzofuran-1(3*H*)-one (benzolactones) 86
4.3.2. Synthesis of the isatin aromatic heads 90
4.4. Coupling of the aromatic heads 94
4.4.1. Coupling with the isobenzofuran-1(3*H*)-one 94
4.4.2. Coupling with the isatin aromatic heads 110
4.5. Concluding remarks 114
CHAPTER 5: Computational and Modelling Studies of the CHIKV nsP2 as a Possible Drug Development Target 117
5.1. Introduction 117
5.2. Results and discussion 120
5.2.1. Protease active site (domain C) 120
5.2.2. N domain binding site 122
5.3. Virtual screening with the CHIKV nsP2 123
5.3.1. Analysis of the C domain virtual screening results 136
5.3.2. Analysis of the N domain virtual screening results 141

5.4. Concluding remarks 144
CHAPTER 6: Structure Based Design towards the Identification of Novel Binding Sites and Inhibitors for the Chikungunya Virus Envelope Proteins 146
6.1. Introduction 146
6.2. Results and discussion 149
6.2.1. Identification of novel binding sites 149
6.2.2. Virtual screening with the CHIKV envelope proteins *150*
6.3 Analysis of the docking calculations 173
6.4. Concluding remarks 180
CHAPTER 7: New Leads for African Trypanosomiasis 181
7.1. Introduction 181
7.2. Results and discussion 184
7.3. Concluding remarks 193
CHAPTER 8: Conclusions and Future Directions 194
8.1. Chapter 2: The Search for anti-CHIKV lead compounds 194
8.2. Chapter 3: The CHIKV nsP3 197
8.3. Chapter 4: Mycophenolic acid analogues 198
8.4. Chapter 5: The CHIKV nsP2 200
8.5. Chapter 6: The CHIKV envelope proteins 201
8.6. Chapter 7: New leads for african trypanosomiasis 203
CHAPTER 9: Experimental 205
9.1. Chemistry 205
9.2. *In silico* experiments 284
9.2.1. CHIKV non-structural protein 2 (nsP2) 284
9.2.2. CHIKV envelope proteins 286
9.2.3. CHIKV non-structural protein 3 (nsP3) 289
9.3. Biological evaluation 290
9.3.1. Anti-Chikungunya evaluations 290
9.3.2. Trypanocidal activity evaluation 290
CHAPTER 10: References 293

Publications

1- **Rashad, A. A.**; Keller, P. A. Structure Based Design towards the Identification of Novel Binding Sites and Inhibitors for the Chikungunya Virus Envelope Proteins. *Journal of Molecular Graphics and Modelling* **2013**, 44, 241-252.

2- **Rashad, A. A.**; Mahalingam, S.; Keller, P. A. Chikungunya Virus: Emerging Targets and New Opportunities for Medicinal Chemistry. *Journal of Medicinal Chemistry* **2014**, 57 (4), 1147–1166.

3- **Rashad, A. A.**; Jones, A. J.; Avery, V. M.; Baell, J.; Keller, P. A Facile Synthesis and Preliminary Structure Activity Analysis of New Sulfonamides Against Trypanosoma brucei. *ACS Medicinal Chemistry Letters* **2014**, 5 (5), 496–500.

Acknowledgements

First and foremost, to Paul Keller, his keen supervision, expertise, encouragement and patience has enriched these years with an extremely astonishing experience. His good care and advice has had great influence on developing me as a scientist. I have gained a fruitful experience in chemistry, team working as well as scientific writing skills based on his directions. This has greatly allowed me to develop my career as an academic and a researcher.

To all members of the Keller research group, I have enjoyed working with you all and we have spent very good times together inside and outside the lab. I have also gained lots of different experiences form the different cultures and expertise backgrounds within the group. Special thanks to Andrew Stevens, my first lab colleague, for our fruitful discussions during my early days. I would like also to thank Mohammed for his guidance in molecular modelling.

Special thanks and gratitude to those who were responsible for the biological evaluation; Prof. Johan Neyts and Dr. Pieter Leyssen, Rega Institute of Medical Research, Belgium, for the anti-chikungunya testing, Prof. Suresh Mahalingam, Emerging Viruses and Inflammation Research group, Institute for Glycomics, Griffith University, for the anti-chikungunya testing and to Dr. Amy Jones, Discovery Biology, Eskitis Institute, Brisbane Innovation Park, for the anti-trypanosoma testing.

Thanks to Dr. Wilford for training on NMR machines and his guidance and fruitful discussions along the way with the NMR problems during this work.

A huge thanks to Dr. Glennys for all of the first year teaching opportunities as well as always checking to see how I was going, discussing the situation in Egypt and asking about my family.

Finally, I would like to express my deepest thanks and gratitude to my family member; my parents, sisters and brother, for their emotional support and encouragement. Huge thanks to my wife, for her support, love and her belief in me, which has greatly pushed me to keep going through this degree.
To everybody in the school of chemistry, thank you very much for the excellent friendly environment.

Abstract

Chapter 2 discusses the synthesis of the arenearylpyrimidylmethanes (AAPMs) series was investigated to further develop the structure activity relationships (SAR) of these compounds as potential anti-chikungunya virus agents. 1-(4,6-Dichloropyrimidin-5-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol was prepared in 50% yield. Subsequent dehydration and amination gave the intermediate 4-amino-6-chloro-5-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyrimidine in 90% yield, over two steps. Final amination using either 1-amino-3-(diethylamino)propan-2-ol or N^1,N^1-diethylpentane-1,4-diamine was attempted using several conditions, however, the desired final AAPM derivatives were not obtained.

Further modification of the synthetic protocol was achieved through the replacement of the 4,6-dichloropyrimidine with the less hindered 4-chloropyridine. The 4-chloropyridin-3-yl alcohols were prepared in 42-52% yield. Subsequent dehydration afforded the 4-chloropyridine alkenes in 62-85% yield. Final replacement of the chlorine atom with the amine N^1,N^1-diethylpentane-1,4-diamine was attempted under different conditions, however, the desired products were not obtained. Instead the 3-(cyclobutylidene(4-phenoxypyridin-3-yl)methyl)aniline was isolated when using phenol as a solvent for the amination reaction. Replacement of the N^1,N^1-diethylpentane-1,4-diamine with the secondary amine morpholine gave 4-(3-(cyclopropylidene(3-nitrophenyl)methyl)pyridin-4-yl)morpholine in 65% yield.

Chapter 3 discusses a development of a virtual screening computer model for the Chikungunya virus (CHIKV) non-structural protein 3 (nsP3) macro domain, using the ADP-ribose binding site as a possible druggable pocket for potential inhibitors. The NCI diversity set III of 1990 compounds were screened against the nsP3 using Autodock Vina as the virtual screening engine with 17 hits initially identified as stronger ligands than the original co-crystallized ligand (ADP-ribose). The hits were further screened in a second round docking using Autodock 4 where 7 hits were ranked higher than ADP-ribose. The N^1,N^4-bis(3-(1*H*-benzo[*d*]imidazol-2-yl)phenyl) terephthalamide hit achieved a binding energy (ΔG_{bind}) of -11.31 Kcal/mol and a predicted inhibitory constant (K_i) value of 5.1 nM and was selected for further *in silico* docking-based optimization. As a result, N^1,N^2-bis(3-(5-cyano-1*H*-benzo[*d*]imidazol-2-

yl)phenyl)oxalamide was selected as an optimized ligand with ΔG_{bind} of -13.12 Kcal/mol and K_i value of 0.239 nM, a more than 21 fold improvment in the combined *in silico* binding profile to the nsP3 active site. This optimized ligand was chemically accessed in three facile synthetic steps, starting with the synthesis of the 2-(3-nitrophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile. Subsequent reduction of the nitro group afforded the 2-(3-aminophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile in 80% yield. Reaction with oxalyl chloride finally afforded the target optimized hit in 69% yield. Compounds are under anti-chikugunya and cytotoxicity evaluation.

In chapter 4, a series of benzolactone-acid conjugates, benzolactone-tetrazole conjugates and isatin-acid conjugates were prepared as mycophenolic acid derivatives to be evaluated as potential inhibitors for the CHIKV. The benzolactones were synthesised through the Pd catalysed cyclization of benzoic acids. The isatin cores were accessed from the corresponding aniline derivatives. Suzuki coupling with 3-(3-boronophenyl)propanoic acid afforded the acid conjugates. The tetrazole conjugates were accessed *via* two pathways: the formation of the acetonitrile intermediates *via* a Suzuki reaction with either 3-(3-boronophenyl)propanoic acid or 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetonitrile, with subsequent cycloaddition reaction with NaN_3. The second pathway to access the tetrazole conjugates was through a one step Suzuki reaction with the synthesized potassium 5-((3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)methyl)tetrazol-1-ide. The synthesized benzolactones-acid conjugates, benzolactone-tetrazole conjugates and isatin-acid conjugates have drug like qualities, as they are water soluble and have $LogD_{5.5}$ values in the range of 1-4. Therefore, this series could be further developed as orally bioavailable compounds. These compounds are under anti-chikugunya evaluation.

In chapter 5, the crystal structure of the CHIKV nsP2 (the viral protease) was investigated for druggable binding sites for the development of possible inhibitors. Two sites were detected, one within the proteolytic C domain, and one within the NTPase/RTPase N domain. A computer virtual screening model was developed for each site, where the Life Chemicals cysteine protease inhibitor library of 28,960 compounds was screened in both sites using the FRED virtual screening engine. The interacting

residues were determined for each hit. A second round of refining docking was performed on the top FRED hits, using Autodock. The top 5 hits in each site were recorded and purchased for anti-chikungunya evaluation.

Chapter 6 discusses the development of a virtual screening computer model to identify novel binding sites for the chikungunya envelope glycoproteins. The model enabled the identification of possible antagonists for these sites through virtual screening using two successive docking scoring functions; FRED docking for fast precise screening, with the top hits then subjected to a ranking scoring using the Autodock algorithm. Both the immature and the mature forms of the chikungunya envelope proteins were included in the study to increase the probability of finding positive and reliable hits. Some small molecules were identified as good *in silico* chikungunya virus envelope proteins inhibitors, representing templates for drug design targeting this virus.

In chapter 7, three facile synthetic steps were developed to synthesize analogues for the lead *N*-(1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2-fluorobenzenesulfonamide that was discovered through a high throughput screening of 87,926 compounds (WEHI 2003) library against *Trypanosoma brucei*. The first step in the synthetic pathway was the sulfonation of the secondary amine of the commercially available 7-nitrotetrahydroquinoline using the appropriate sulfonyl chlorides. Subsequent reduction using Raney Nickel and hydrazine hydrate afforded the 7-aminotetrahydroquinoline intermediates which were reacted with the appropriate sulfonyl chlorides to afford the final bis-sulfonamide analogues. Evaluation of the anti-trypanosoma activity of the compounds revealed the activity of some analogues with the IC_{50} values in the range of 2-4 μM, with good selectivity indices.

List of Abbreviations

°C	Degree(s) Celsius
^{1}H NMR	Proton Nuclear Magnetic Resonance
^{13}C NMR	Carbon Nuclear Magnetic Resonance
2D	Two Dimensional
3D	Three Dimensional
AAPM	Arenearylpyrimidylmethane
ADME	Absorption, Distribution, Metabolism and Elimination
ADT	Autodock tools
Ala	Alanine
Ar	Aromatic Ring
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
br	Broad
C	Capsid Protein
cal	Calories
CHIKF	Chikungunya fever
CHIKV	Chikungunya virus
cLogD	Calculated LogD
clogP	Calculated logP
Cys	Cysteine
d	Doublet
DCB	1,2-dichlorobenzene
dd	Doublet of Doublets
ddd	Doublet of Doublet of Doublets
DIPA	Diisopropyl Amine
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
dt	Doublet of Triplets
DV	Dengue Virus

EC_{50}	Concentration to achieve 50% Induction of the response
EI-MS	Electron Impact Mass Spectrometry
eq.	Equivalent(s)
ER	Endoplasmic Reticulum
ESI-MS	Electrospray Ionisation Mass Spectrometry
FDA	Food & Drug Administration
FRED	Fast Rigid Exhaustive Docking
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
GMP	Guanosine Mono Phosphate
GTP	Guanosine 5′-Triphosphate
GUI	Graphics User Interface
h	Hour(s)
HAT	Human African Trypanosomiasis
Hb	Hydrogen Bond
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HCV	Hepatitis C Virus
His	Histidine
HIV	Human Immunodeficiency Virus
HIV RT	Human Immunodeficiency Virus Reverse Transcriptase
HPLC	High Performance Liquid Chromatography
HR	High Resolution
HTS	High Throughput Screening
Hz	Hertz
IC_{50}	Concentration to Effect 50% Inhibition
IFN	Interferon
Ile	Isoleucine
IMPDH	Inosine Monophosphate Dehydrogenase
K	Kilo
LDA	Lithium Diisopropylamide
Leu	Leucine

LR	Low Resolution
Lys	Lysine
m	Multiplet
m.p.	Melting Point
Met	Methionine
min	Minute(s)
mol	Mole(s)
MPA	Mycophenolic acid
Mse	Selenomethionine
Mwt	Molecular Weight
NCI	National Cancer Institute
nM	NanoMolar
NMR	Nuclear Magnetic Resonance
ns	Non-structural
NTPase	Nucleotide Triphosphatase
ORFs	Open Reading Frames
p	Pentet
P. falciparum	*Plasmodium falciparum*
PHB	Prohibitin
Phe	Phenyl Alanine
PLC	Preparative Thin Layer Chromatography
Pro	Proline
q	Quartet
RNA	Ribonucleic Acid
RNAi	Ribonucleic Acid Interference
rt	Room Temperature
RTP	Ribavirin Triphosphate
s	Singlet
SAR	Structure Activity Relationship
Ser	Serine
sp	Septet
sex	Sextet
SFV	Semliki Forest Virus

SINV	Sindbis Virus
t	Triplet
td	Triplet of Doublets
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
TFA	Trifluoroacetic acid
Thr	Threonine
Trp	Tryptophan
tt	Triplet of Triplets
Tyr	Tyrosine
UV	Ultraviolet
Val	Valine
w	Weak
W	Watts
WNV	West Nile Virus

CHAPTER 1: Introduction

1.1. **Classification, History and Clinical Features**

Chikungunya virus (CHIKV) is an emerging arthrogenic arbovirus that belongs to the *alphavirus* genus, family *Togaviridae*. It has been responsible for major outbreaks of a devastating human arthritis disease during the past five years.[1] Chikungunya fever caused by the virus was first described in 1952,[2] after an outbreak on the Makonde Plateau (named after an ethnic group from East Africa), along the border between Tanganyika and Mozambique. During this period, a high proportion of residents of all ages were affected by a distinctive disease with a sharp onset of crippling joint pains, severe fever, and a conspicuous rash.[2] The elders of the Makonde tribes could not remember any previous, similar epidemics with these symptoms, suggesting that this was a new illness. The word ''Chikungunya'' translates to ''that which bends up'' relating to the stooped posture developed as a result of rheumatologic inflammation.[3] After that, only minor outbreaks occurred periodically in Africa, but some major epidemics were reported in the 1960s and 70s in India and Southeast Asia.[4] After the 1973 outbreak in India, only sporadic activities were detected over the next 30 years, with no major recurrence until a large outbreak in Kenya in 2004.[5] This outbreak initiated a spreading epidemic that reached numerous islands of the Indian Ocean, India, and parts of Southeast Asia, and was also detected in 18 countries throughout Asia, Europe, and North America *via* imported infectious carriers. Over the course of five years, an estimate of more than 2 million cases occurred, with outbreaks in several countries where the virus had not been previously documented.[6] The first European detection of CHIKV occurred in Italy in 2007.[7,8]

Nearly 40 countries have detected chikungunya virus infected cases so far (Figure 1.1). The virus was listed in 2008 as a category C priority pathogen by the US National Institute of Allergy and Infectious Diseases (NIAID) that include pathogens which could be engineered for mass dissemination in the future, because of the high morbidity and mortality rates and major health impact.[9,10] Recent epidemics were reported in India (1.4 to 6.5 million cases in 2006-2007), 3,000 – 42,000 cases were detected in 2009 in Malaysia and Thailand.[11,12] The CHIKV mortality rate has been

estimated to be 1 in 1000 and most of the deaths occur in neonates, adults with underlying conditions and the elderly.[10]

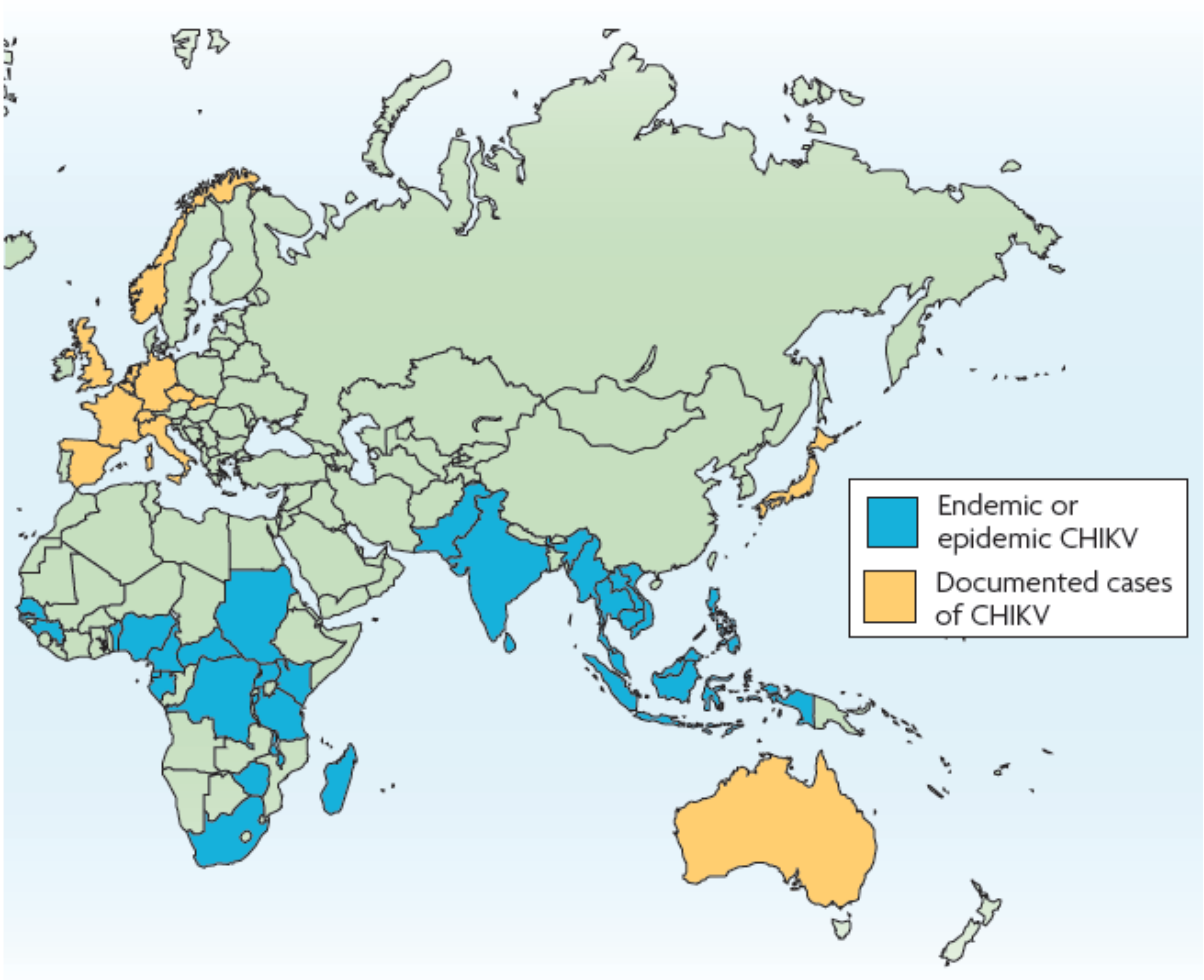


Figure 1.1 Chikungunya virus Worldwide distribution,[10] figure used with permission from NPG (Nature publishing group).

CHIKV can be transmitted through an urban cycle, man to mosquito to man, or a sylvatic cycle, animal to mosquito to man.[13] The virus is transmitted to humans by mosquitoes of the *Aedes* genus (*Aedes furcifer* in Africa and *Aedes aegypti* in Asia), similar to the dengue fever causing virus. However, the *Ae. albopictus* mosquito was extensively implicated in CHIKV transmission during the 2005-2006 outbreak on Reunion island.[10]

The switch of the CHIKV vector from *Ae. aegypti* to *Ae. albopictus* was caused by the insufficient *Ae. aegypti* number for its transmission.[14] A mutation in the E1 envelope protein, A226V, increased the CHIKV fitness in *Ae. albopictus* and improved the transmissibility of the virus through *Ae. albopictus* to vertebrate species.[15] *Ae.*

albopictus has spread to Madagascar, the Indian Ocean nations, Africa, Southern Europe and the USA.[16,17] Virus transmission has been also reported as a result of maternal-foetal transmission in recent epidemics.[18]

Chikungunya fever *(CHIKF)*: The symptoms of CHIKF infection generally start 4–7 days after the mosquito bite. Infection tends to present in two phases, the first stage is acute, while the second stage is persistent (chronic), causing disabling polyarthritis.[19] Acute infection lasts 1–10 days and is characterized by a painful polyarthralgia, high fever, asthenia (weakness), headache, vomiting, rash, and myalgia (muscle pain). Rash is the least reliable symptom, presenting in as few as 19% of patients. When a rash is present, it is typically maculopapular in nature, but recent studies have also noted vesiculobullous lesions with desquamation.[20] The persistent chronic CHIKF phase is characterized by polyarthralgia (aches in the joints, joint pains) that can last from weeks to years beyond the acute stage.[21] Eighteen months after disease onset, 40% of patients are found to still have anti-CHIKV immunoglobulin M (IgM).[22]

CHIKV attacks fibroblasts, explaining the involvement of muscles, joints, and skin connective tissues. The high number of nerve endings within the joints and muscle tissues explains the pain associated with CHIKF. Neurological manifestations have also been reported during the most recent epidemics. In India, some neurological disorders such as encephalitis, peripheral neuropathy, myelopathy, myeloneuropathy and myopathy were reported.[23] Moreover some cases with multi-organ failure have also been reported.[24] Eye infection (Chikungunya neuroretinitis) has also been reported where patients suffered from a sudden, painless diminution of vision in both eyes.[25,26] It has been also reported that CHIKV can infect the cornea and can be transmitted *via* the ocular route.[27]

1.2. Virology of the CHIKV

The CHIKV genome (Figure 1.2) is a positive sense, single stranded RNA genome of about 11.8 Kb in size. It consists of two open reading frames (ORFs),[28] one in the 5` end encoding the non-structural protein (nsP) precursors:

- nsP1: involved in viral mRNA capping *via* its guanine-7-methyltransferase and guanylyltransferase enzymatic activities,

- nsP2: acts as protease and helicase,
- nsP3: part of the replicase unit and an accessory protein involved in RNA synthesis,
- nsP4: RNA-dependent-RNA polymerase,

The nsP123 precursor and nsP4 function in a complex for viral negative-strand RNA synthesis. The 3` end ORF encodes the structural proteins, the capsid (C), envelope glycoproteins E1 and E2 and two small cleavage products (E3, 6K). The untranslated junction region (J) (Figure 1.2) contains its internal promoter, a conserved sequence of 21 nucleotides, for transcription of the sub-genomic mRNA in other *alphaviruses* (sindbis virus).[29]

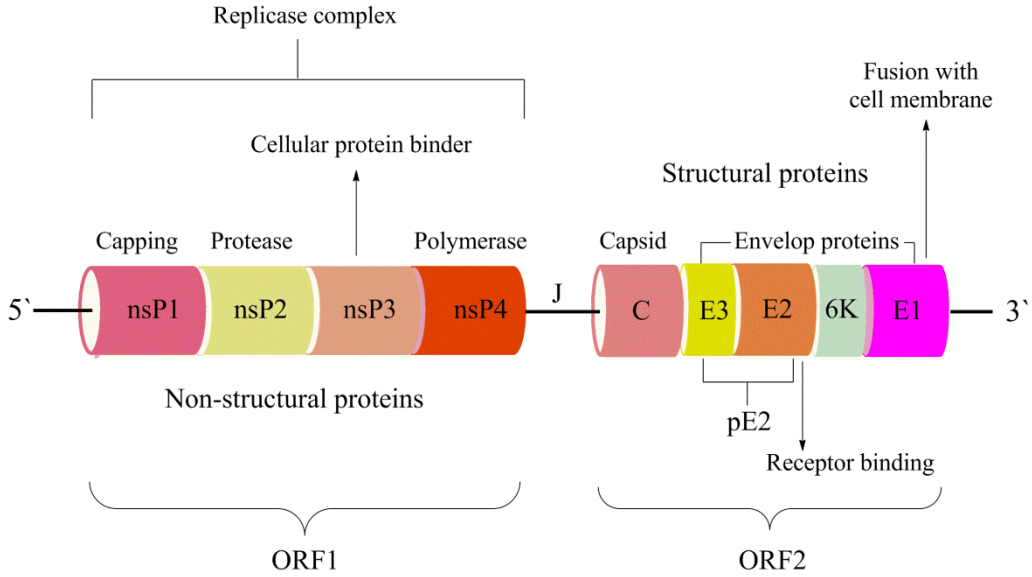


Figure 1.2 Schematic representation for The CHIKV genome showing the RNA sequence ORFs.

The CHIKV surface consists of 80 trimeric spikes composed of heterodimers of the envelope glycoproteins (E1 and E2) in the lipid bilayer. Similar to other members of the *alphaviruses*, the CHIKV starts its life cycle (Figure 1.3) by entering the target cells by pH dependent endocytosis in clathrin coated vesicles *via* receptor mediated interaction,[30] but the exact mechanism by which it does so remains unclear. CHIKV has been shown to replicate in a large number of cell types including epithelial, endothelial and fibroblast cells as well as monocyte derived macrophages.[30] A recent study identified prohibitin (PHB) as a microglial cell expressed CHIKV binding protein.[31] PHB is an evolutionarily conserved and ubiquitous protein that consists of two highly homologous proteins of different molecular weights. PHB1 has a mass of approximately

30 kDa while PHB2 is approximately 37 kDa. The two proteins oligomerize, and hetero-oligomerization is essential for protein stability.[32] PHB has been shown to be present in multiple cell compartments including the mitochondria, cytoplasm and nucleus in addition to its expression on the cell surface.[33] PHB1 was confirmed as a CHIKV E2 binding protein, but not PHB2.[31] PHB1 was found to be involved in the internalization process either on its own or as part of a complex, further suggesting that a PHB-virus interaction may be mediated by the specific PHB molecule that interacts with the virus. Experimental down-regulation of PHB1 significantly reduced the level of infection in tested cell lines. However, it is believed that this mechanism might be only one pathway by which CHIKV can enter the susceptible cells, as in case of other enveloped viruses.[31]

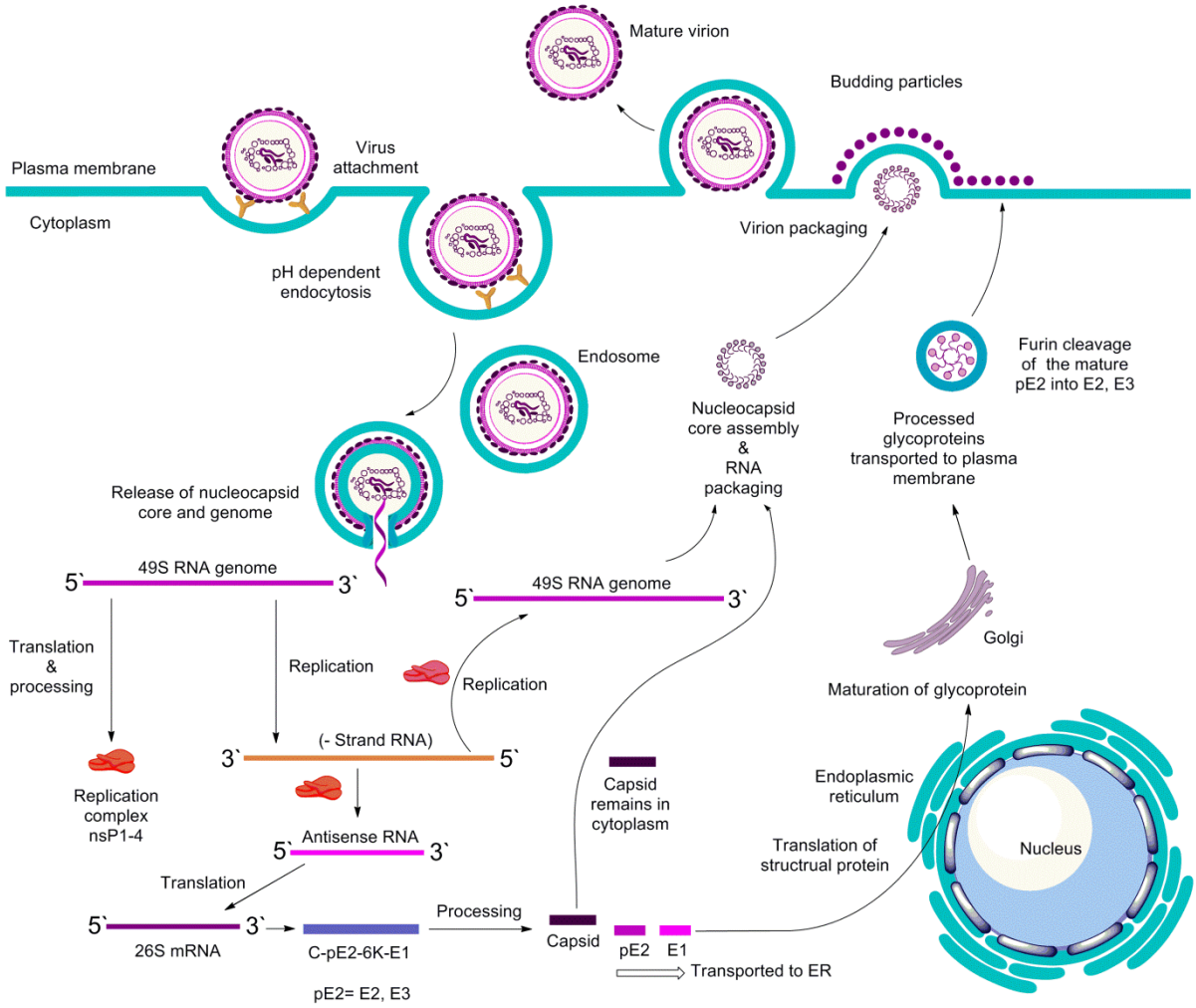


Figure 1.3 Schematic representation of the Chikungunya virus life cycle.

After entering the cell, the endosome acidic environment triggers conformational changes, the viral envelope proteins (E1 and E2 complex) reorganizes, leading to

dissociation of the E2-E1 heterodimers, and the formation of E1 homotrimers. E1 trimerizes and inserts into the target membrane with the hydrophobic fusion peptide (fusion loop) and refolds to form a hairpin-like structure. Exposure of the E1 fusion peptide mediates virus host cell membrane fusion, releasing the nucleocapsid into the cytoplasm.[34,35] This process depends on low pH and cholesterol, which is also required for budding during *alphavirus* infection.[36,37]

Two non-structural protein precursors are translated from the viral mRNA, and then are cleaved generating nsP1, 2, 3 and 4. During translation, nsP123 binds to free nsP4 and with some cell proteins, forming the replication complex,[38,39] which synthesizes a full-length negative-strand RNA intermediate required for replication. When nsP123 concentration increases, it is cleaved into nsP1, nsP2, nsP3 and nsP4 which forms, along with host cell proteins, the positive strand replicase, which produces the 26S sub-genomic positive strand RNAs and genomic (49S) RNAs.[39] Promoters present in the negative strand initiate the transcription of 26S sub-genomic positive stranded RNA which encodes the structural protein precursors. The later is cleaved by a serine protease to yield the capsid (C) which remains in the cytoplasm (Figure 1.3), pE2, 6K and E1.[40] The C protein might be responsible for such autoproteolytic activity as it has few conserved sequences which have similar activity, and are common in other *alphaviruses*.[41]

pE2 and E1 are translated in the endoplasmic reticulum (Figure 1.3) and processed in the Golgi, and are moved to the plasma membrane, where pE2 is cleaved by furin-like protease activity in the host cell into E2 and E3.[42] The assembly of virions begins in the cytoplasm of the cell, where the formation of the nucleocapsid with 120 dimers of the C protein starts to occur.[43] The assembled particle buds at the cell membrane as spherical particles of 65 to 70 nm in diameter, composed of genomic RNA molecules and the capsid proteins and enveloped in a host-derived lipid membrane.

1.3. The Development of CHIKV Vaccine

There is an urgent need to control the spreading CHIKV, however, there is little understanding of the interaction between the chronic CHIKV infection and the immune system in defending the body against any subsequent reinfection.[10] The immune responses are strongly accepted to induce autoimmunity, by cross reactivity between

viral and host antigens. The B cells and T cells might respond to CHIKV and this may contribute to the long-term joint disease experienced by many convalescent patients.[44]

Thus far, there is no licensed CHIKV vaccine. Some vaccine preparations that involved either formalin inactivation or Tween-ether extracts of virus[45] showed high immune responses without any adverse effects. In 2000, the US Army carried out a Phase II clinical trial examining a live attenuated CHIKV vaccine.[46,47,48] The vaccine was formulated from human MRC-5 cell line. In this study, subjects that received the vaccine developed neutralizing antibodies, with few subjects showed mild joint pain compared to those who did not.[48]

In 2009, one study reported a vero cell adapted formalin inactivated prototype vaccine with alhydrogel as adjuvant that was prepared using an Indian CHIKV strain implicated in the 2006 epidemic. The humoral immune response was characterized by high titer antibodies that have been confirmed through microcytotoxicity assay and *in vivo* neutralization tests. Therefore, this could be a promising, safe and effective vaccine that can elicit long lasting protective immune response.[49]

A live CHIKV vaccine was developed in 2011 that elicits a protective immune response with no detectable disease in mice. It is also unable to infect mosquito vectors, which is an important safety feature for a live virus vaccine that could be used in non endemic areas to immunize travellers or laboratory personnel. However, this vaccine candidate is still under evaluation in nonhuman primates, before evaluation in humans.[50] Other promising CHIKV vaccine candidates that depend on virus-like particles are in early stages of preclinical development.[51,52] A successful virus-like particle vaccine based on viral structural proteins was tested on nonhuman primates and was found to produce neutralizing antibodies that protect against viremia after high-dose challenge. When these antibodies were transferred into an immunodeficient host (a mouse), the host was protected indicating a passive immunity.[53]

With the ongoing vaccine development research against the CHIKV, the world remains under the threat of rapidly spreading CHIKV infections, emphasizing the importance of developing chemotherapeutics targeting the virus.

1.4. Emerging novel CHIKV targets

As previously mentioned, the CHIKV genome is formed from 2 ORFs, one in the 5` end coding for nsP1, nsP2, nsP3 and nsP4. The 3` end ORF encodes capsid (C), envelope

proteins E1, E2, E3 and 6k (Figure 1.2). These proteins, which mediate essential steps in the lifecycle (Figure 1.3) of the virus,[10] could be possible targets for drug design.

1.4.1. Non-Structural Proteins

1.4.1.1. Non-structural protein 1

Like in other *alphaviruses*, CHIKV nsP1 is a palmitoylated 535 amino acid protein. The N-terminal region is a methyltransferase and guanylyltransferase involved in capping and methylation of the newly formed viral genomic and subgenomic RNAs.[54] In early 2013, it was shown that CHIKV nsP1 acts as antagonist for the bone marrow stromal antigen 2 (BST-2).[55] BST-2 is one of the host cell defensive mechanisms, and is induced by interferon (INFα). BST-2 expression results in retaining viruses at the surface of the infected cells.[56] BST-2 was found to co-localize with CHIKV E1 and nsP1, but only nsP1 is able to down-regulate BST-2 expression and therefore, inhibiting the virus tethering on cell surface.[55] This activity of the CHIKV nsP1 is similar to that of the HIV-1 Vpu protein[56] in that both repress BST-2. This discovery will help in developing BST-2 mediated therapeutics targeting the nsP1.

1.4.1.2. Non-structural protein 2

The non-structural protein 2 (nsP2) of *alphaviruses* is a multifunctional protein.[57,58,59,60] The proteolytic domain has been allocated to its C-terminal section which forms a papain like cysteine protease (also known as thiol protease).[57,61] The nsP2 proteolytic activity is critical for virus replication and is responsible for cleavage of the non-structural polyprotein complex.[62,63]

The proteolytic activity of the CHIKV nsP2 has been demonstrated,[64] and the enzymatic activities within the N-terminus have been investigated recently. It was found to have RNA triphosphatase activity that performs the first of the viral RNA capping reactions. It was also found to have a nucleotide triphosphatase (NTPase) activity, fueling the RNA helicase activity performed by the C-terminal domain.[65] CHIKV-nsP2 also has 5`-triphosphatase (RTPase) activity that removes the γ-phosphate from the 5` end of RNA. Both NTPase and RTPase activities are completely dependent on Mg^{2+} ions.[65]

Both N and C domains are composed of α-helices and β-strands (Figure 1.4). The N terminus is dominated by α-helices, whereas the C-terminal domain contains

helices and strands. The central β-sheets are flanked by α-helices. The crystal structure of CHIKV nsP2 protease has been solved and comprises 324 residues. Being a cysteine protease, the catalytic mechanism involves a nucleophilic cysteine thiol in a catalytic dyad.[66] Analysis of the CHIKV nsP2 crystal structure shows 6 cysteine residues, three in the N-terminus (Cys1013, Cys1057 and Cys1121) and three in the C-terminus (Cys1233, Cys1274 and Cys1290) as shown in Figure 1.4. Since the proteolytic activity is isolated in the C-terminus,[64] one of the three cysteine residues in the C-domain might contribute as the catalytic thiol.

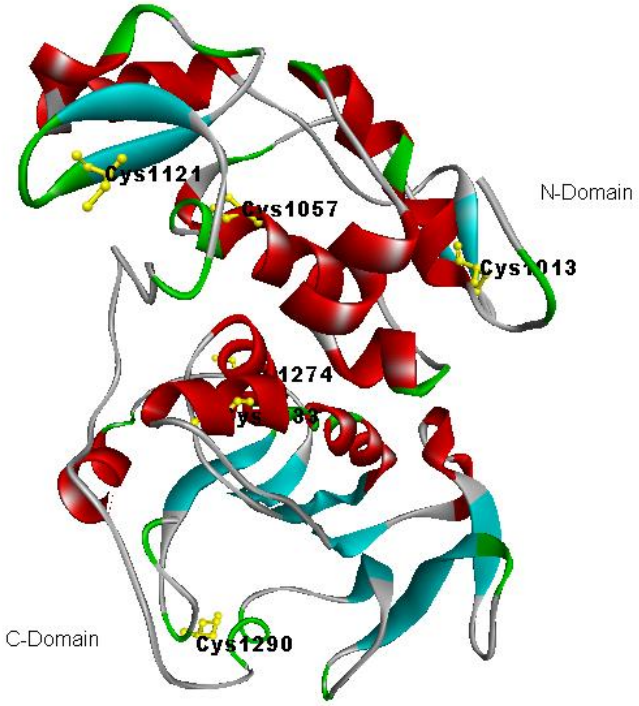


Figure 1.4 CHIKV nsP2 crystal structure showing the N- and C-terminal domains, cysteine residues shown in yellow balls and stick form (PDB code: 3TRK[67], no citation was found for the crystal structure).

The first step in the mechanism of cysteine proteases catalysis is usually the deprotonation of a thiol group within the enzyme active site by an adjacent amino acid containing a basic side chain, often a histidine residue.[68] Among the three cysteine residues in the C-terminus (Figure 1.5), the Cys1274 residue is less likely to be involved in the catalytic mechanism because only one His residue (His1314) is nearby, whereas the other cysteine residues, four His residues, His1222, His1228, His1229 and His1236 could be associated in the deprotonation mechanism (Figure 1.5).

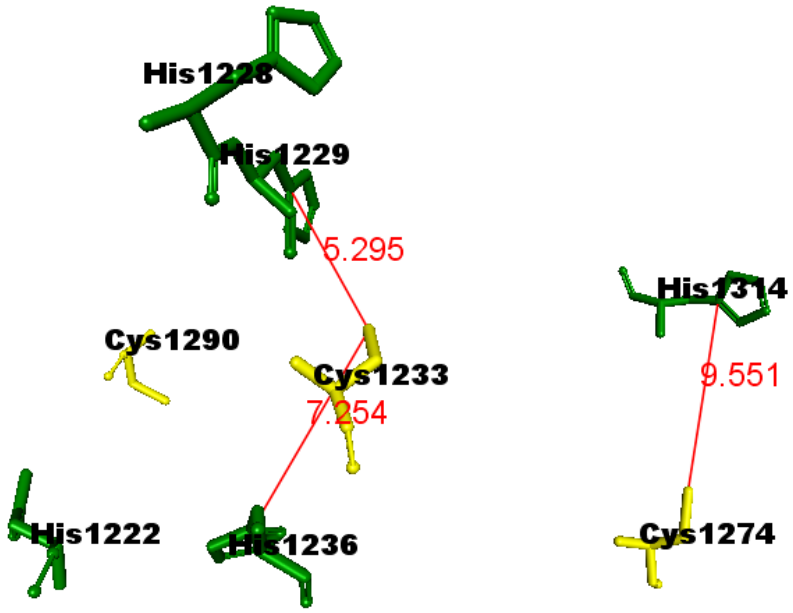


Figure 1.5 Schematic representation of the CHIKV nsP2 C-Domain showing the positions of the cysteine residues (yellow) and histidine residues (green). Distances (Å) shown in red, generated from the crystal structure PDB file code: 3TRK.

In 2012, Singh Kh *et al*. reported the development of a homology model of CHIKV nsP2 protein based on the crystal structure of the nsP2 protein of Venezuelan equine encephalitis virus (VEEV),[69] in order to locate the active site of the protease. The critical residues in nsP2 were identified by docking three different peptides in order to identify the residues responsible for non-structural protein cleavage, nsP1-2, nsP2-3 and nsP3-4 peptides. These three peptide sequences represent the substrates for the nsP2 proteolytic processing with a remarkable preference of nsP3-4>nsP1-2>nsP2-3.[70]

The active site was investigated and was found to lie in the C-terminal domain[69] (Figure 1.5). The key residues Gln1039, Lys1045, Glu1157, Gly1176, His1222, Lys1239, Ser1293, Glu1296 and Met1297 were found to interact with the nonstructural protein sequence complex to be cleaved, and were considered an individual functional unit. Only two residues are located in the N-domain, Gln1039 and Lys1045, all the other residues are located in the C-domain. Similar work by Bassetto *et al.*[71] reported the development of a homology model for the nsP2 protease active site within the C-domain.

The predicted active site by Singh Kh *et al.*[69] and Bassetto *et al.*[71] matches with the abovementioned explanation concerning the poisoning the active site within the C-

domain, especially, they found that His1222 residue to be lying within the predicted active site pocket. Analysis of the enzyme surface shows that the predicted active site is located in a major surface groove as shown in Figure 1.6, the major cavity on the enzyme surface is more likely to accommodate the substrate polyprotein sequence to be cleaved.

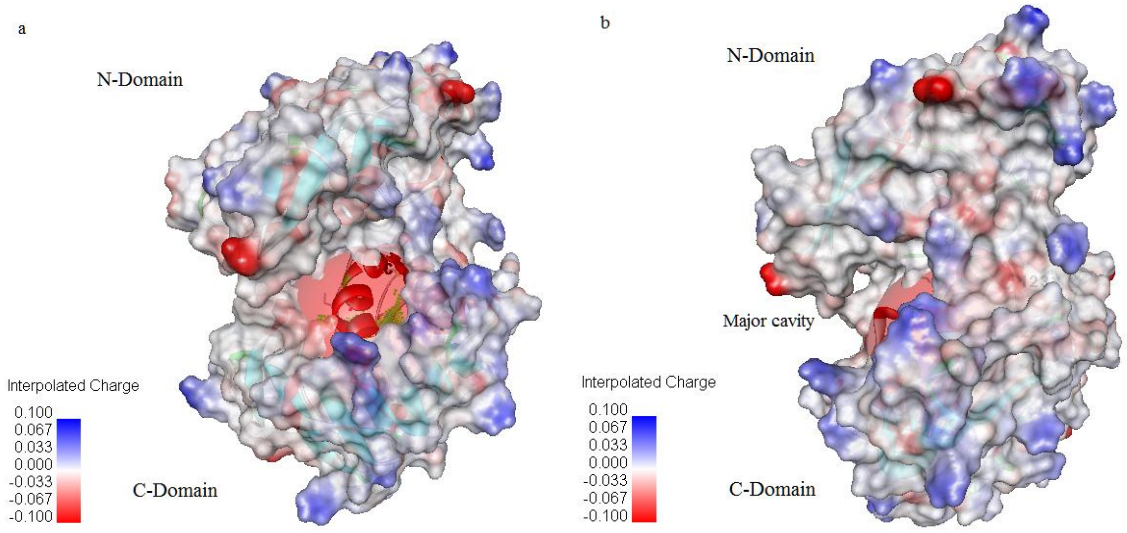


Figure 1.6 Electrostatic potential surface of the CHIKV nsP2 and its active site pocket within the C-domain, a) The active enzyme surface front view showing the active site as a transparent red sphere surrounding the active site with the Cys1233 and His1222 residues lying within the sphere. b) Side view of the enzyme (rotated left by 90º about the vertical axis from the view in a) showing the major accessible cavity to the active site with a part of the sphere protruding toward the outer surface, generated from the PDB file 3TRK.

This major enzyme groove (Figure 1.6b) may act as the enzyme mouth holding the protein to be processed. Therefore, targeting the residues specified above,[69] as well as the residues within the active site, would be an applicable strategy to inhibit the enzyme function and consequently inhibiting the virus replication.

Moreover, the *alphaviruses* nsP2 proteins have been described as virulence factors responsible for the transcriptional and translational shutoff in infected host cells and the inhibition of interferon (IFN)-mediated antiviral responses contributing to the controlling of translational machinery by viral factors.[72,73] This controlling comes through interactions with cellular RNA binding proteins, including heterogeneous nuclear ribonucleoproteins (hnRNPs), ribosomal protein S6 (RpS6), and cellular

filament components. Recently, 22 cellular components were hypothesized to interact with nsP2 or nsP4 contributing to the CHIKV replication, mainly, heterogeneous nuclear ribonucleoprotein K (hnRNP-K) and ubiquilin 4 (UBQLN4). Also, it was noted that the interaction of nsP2 with the tetratricopeptide repeat protein 7B (TTC7B) plays a significant role in the cellular machinery control induced by the CHIKV infection.[74]

1.4.1.3. Non-structural protein 3

The function of *alphaviruses* nsP3 has remained unknown, although mutations can affect different steps of the viral replication machinery.[75] It is constructed of two domains, the first is a unique macro domain in the conserved N-terminal region. The C-terminal region is less conserved and is phosphorylated in about 16 positions on serines and threonines.[76,77] The function of phosphorylation is not understood, but it was found that deletion of these phosphorylated residues decreases the RNA synthesis level.[78] Interestingly, viral pathogenicity of Semliki Forest virus (SFV), another *alphavirus*, is decreased in the absence of phosphorylation on nsP3, and the absence of the C terminus alters SFV neurovirulence.[79]

The N-terminus of nsP3 contains a macro domain (known also as X domain), which binds to ADP-ribose derivatives and RNA, and is able to hydrolyse ADP-ribose-1``- phosphate,[80,81] a side product of cellular pre-tRNA splicing. Therefore, it is believed to control the metabolism of ADP-ribose 1``- phosphate and/or other ADP-ribose derivatives which have regulatory functions in the cell. The ADP ribose–binding site within the nsP3 macro domain is solvent-exposed and points away from the other domains in the nsP23 polyprotein. Based on sequence conservation in *alphaviruses*, it has been shown that residues just after the nsP3 macro domain play a role in positioning of the nsP23 complex cleavage site.[82] It can be inferred from the crystal structure of the nsP23 precursor protein of the closely related *alphaviruses*, SINV, that the nsP2 is connected to the nsP3 through the macro domain of the nsP3.[83] The nsP23 cleavage site is located in a narrow cleft formed between nsP2 and nsP3 that is inaccessible for proteolysis, and all the nsP2 noncytopathic mutants lie at the interface between nsP2 and nsP3.[83] The inaccessibility of the nsP23 cleavage site indicates that access is tightly regulated. It is believed that the activator segment is located in the amino-terminus of the nsP2 which becomes exposed after cleavage from the nsP12 precursor poly protein.[84]

In 2010, the crystal structure of the nsP3 macro domain for the CHIKV was solved[80] (Figure 1.7). It is formed of 672 residues and contains six-stranded β sheets with three α helices. The intermolecular interactions between the residues in the binding pocket of the enzyme and the ADP-ribose,[85] as analysed from the crystal structure, are shown in Figure 1.8.

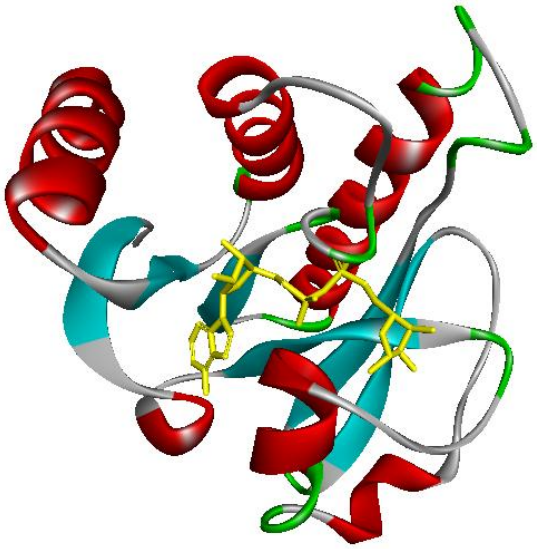


Figure 1.7 Crystal structure of the CHIKV macro domain with the bound ADP-ribose (yellow colour), generated from the PDB file code: 3GPO.

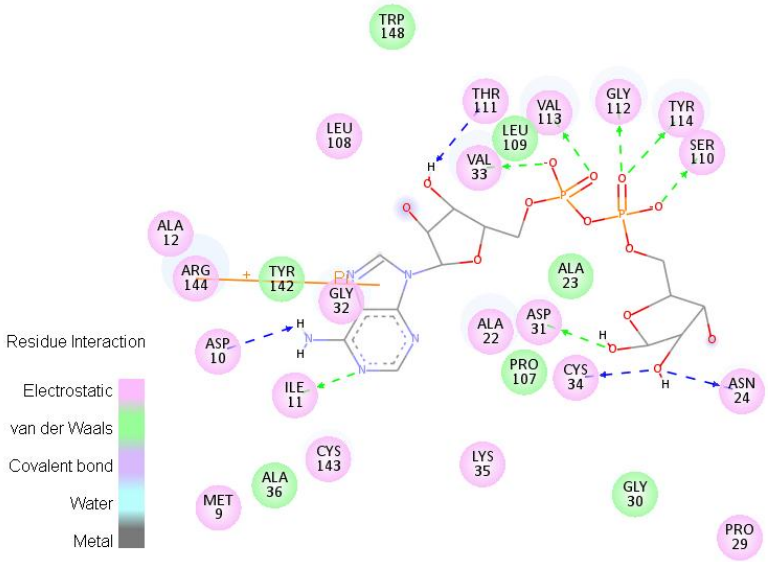


Figure 1.8 2D representation of the interaction of ADP-ribose inside the nsP3 macro domain binding pocket showing the other residues inside the active site (Generated from the PDB file: 3GPO)

As shown in Figure 1.8, the key binding residues are: Arg144, Asp10, Ile11 Thr111, Gly112, Ser110, Tyr114, Val113, Asn24, Asp31 and Val33. The binding

complex is formed from 10 H-bonds and one π-cation interaction.[85] The PO_4^{2-} moiety showed the strongest interactions with these residues in the enzyme pocket. Also, the ribose (with Thr111) and the diphosphate (with Val 33, Ser110, Gly112, Val113, Tyr114) units were found to play major roles in the CHIKV nsP3 ADP-ribose complex.[85]

Understanding this binding interaction of the ADP-ribose to the macro domain of the CHIKV could therefore be a useful element to further assist in drug design and development of inhibitors for this virus. Bound inhibitors to the ADP-ribose binding pocket will alter the function of the nsP3 either cleaved or in a polyprotein complex, and consequently will alter the function of the viral replication machinery.

Recently, more insights on the molecular function of the nsP3 revealed interesting findings, it was reported that the nsP3 of CHKV, as with other *alphaviruses*, use a conserved proline-rich motif to interact with the Src-homology-3 (SH3) domain of amphiphysin-1 and amphiphysin-2 proteins of the host cell, two related member proteins of the BAR (Bin-Amphiphysin-Rvsp) protein superfamily implicated in several cellular functions.[86]

More recently,[87] the nsP3 has been shown to be the inhibitor of stress granule assembly by recruiting G3BP into cytoplasmic foci. The conserved nsP3 SH3 domain-binding motif (the proline-rich motif) is essential for both nsP3-G3BP interactions and viral RNA replication. G3BP (Ras GTPase-activating protein-binding protein) is an enzyme in human cells and a member of the heterogeneous nuclear RNA-binding proteins.[88] This protein plays a major role during infection and in the assembly of stress granules. Stress granules are membranous cytoplasmic focal structures (foci) that immediately aggregate in response to cellular stress, this last action leads to impaired translation of most mRNAs.[89] These stress granules may have antiviral activity that is inhibited by CHIKV replication by the nsP3 SH3 domain-binding motif.[87]

1.4.1.4. Non-structural protein 4

The non-structural protein 4 was identified as the RNA-dependent-RNA polymerase,[29,90] and recently was found to suppress the host cell unfold protein response (UPR), also named as endoplasmic reticulum (ER) stress response.[91] The UPR is a mechanism that maintains the cellular protein homeostasis and prevents over-loading of unfolded protein in the lumen of the ER during normal and diseased cellular conditions.

The UPR involves some steps and some contributing proteins, including the PKR-like ER kinases (PERK).[92] During the UPR, PERK activates by self-dimerization and phosphorylation; this activated PERK in turn, phosphorylates the serine 51 position of the eukaryotic translation initiation factor 2, alpha subunit (eIF2α), an essential factor for protein synthesis. Phosphorylated eIF2α inhibits the general protein synthesis, and consequently, will inhibit the pathogen protein replication. The CHIKV nsP4 was found experimentally to significantly reduce the phosphorylation (serine 51) of eIF2α, and thus ensuring the translation of the viral protein.[91] This discovery can be exploited as a possible target for anti-CHIKV intervention.

1.4.2. Structural Proteins

The invasion of susceptible cells by the CHIKV is performed by two viral glycoproteins, E1 and E2. Both carry the basic antigenic determinants and form the icosahedral shell of the virion particle. E2 and E3 are produced from furin cleavage of p62 (also known as PE2, see Figure 1.2) precursor. E2 is responsible for receptor binding whereas E1 mediates the membrane fusion.[10] E3 contains the 64-amino-terminal residues of p62.[93]

E1 and p62 peptide are type I membrane proteins and are derived from a structural polyprotein precursor. They are translated in the infected cell endoplasmic reticulum, into a p62-E1 heterodimer and processed by the Golgi (Figure 1.3), E3 protects the E2-E1 heterodimer from premature fusion with cellular membranes.[94] The heterodimers trimerize forming the viral spikes. Cleavage of p62 into E3 and E2 during transport to the cell surface prepares the spikes for the fusogenic activation to enter the cell. At the plasma membrane, the formed virions bud through interactions between E2 and genome-containing viral nucleocapsids in the cytoplasm.[95]

In a recent study, the roles of four amino acid residues (G91, V178, A226, and H230) in the CHIKV E1 protein have been linked to the E1 and cell fusion process.[96] The study revealed that the highly conserved amino acid residues, G91 and H230, are important for membrane fusion functionality. The glycine residue (G91) is critical for the fusion process whereas any mutation or substitution in this residue leads to complete loss of E1 fusion ability. The E1 histidine 230 is located outside of the fusion sequence, but is still critical for the fusion. Other structural proteins also affect the E1 fusogenic capacity. E2 protein facilitates both E1 folding and regulates E1 fusogenic properties.

This process is pH and cholesterol dependent.[96] As an *alphavirus* family member, the hydrophobic fusion peptide of the CHIKV was found to be a trimer of hairpins composed of *β*-sheets in the post fusion state (type II fusion proteins).[97,98] Figure 1.9 shows the crystal structure of the CHIKV fusion peptide,[99] consisting of 18 amino acid residues, residues 84−101 in the full-length E1 glycoprotein.[95]

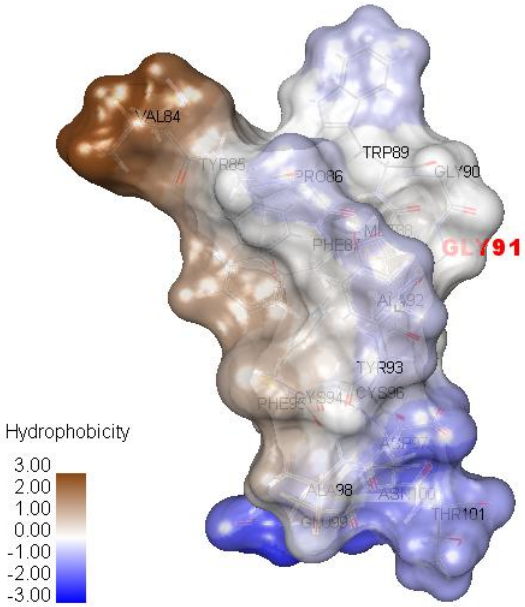


Figure 1.9 Hydrophobic surface view of the CHIKV fusion peptide showing the residue G91 in red label located at the back side (generated from the crystal structure of the fusion peptide, PDB file code: 2RSW[99]).

The crystal structures of the CHIKV p62-E1 (immature) and the E3-E2-E1 (mature) glycoprotein complexes have been recently solved (Figures 1.10-1.12). E1 is folded into three domains I, II and III (Figure 1.11) that are rich in β-sheets. E2 is an immunoglobulin β protein, with three domains (A, B and C). Domain A is at the centre, domain B is at the membrane upper end whereas domain C is towards the viral membrane. The later binds to domain II of E1 by hydrogen bonding due to the hydrophilic contact area between them (Figure 1.10). The long β – ribbon of E2 makes most of the connection with E3 (Figure 1.12). Furin loop (Figure 1.10) is the E2E3 junction in the immature complex, and this junction contains a functional proprotein convertase motif which is cleaved by the cellular proteases, furin-like proprotein convertases, during the maturation (Figure 1.3) of the glycoproteins.[42] Variations within this junction site among the different CHIKV isolates greatly affect the cleavage

susceptibility by furin proteases. The amino acid His60 (residue 56 in the crystal structure PDB file: 3N40) is the critical residue that determines the spectrum of furin and furin-like convertases that process the E2E3 glycoprotein complex.[100] It can be inferred from the comparison between the structures of the immature and the mature glycoprotein complex,[95] that the short peptide sequence (Pro59, His60, Arg63, Glu64, Ser65, Thr66, Lys67 and Asp68) is cleaved form the immature complex after furin cleavage.

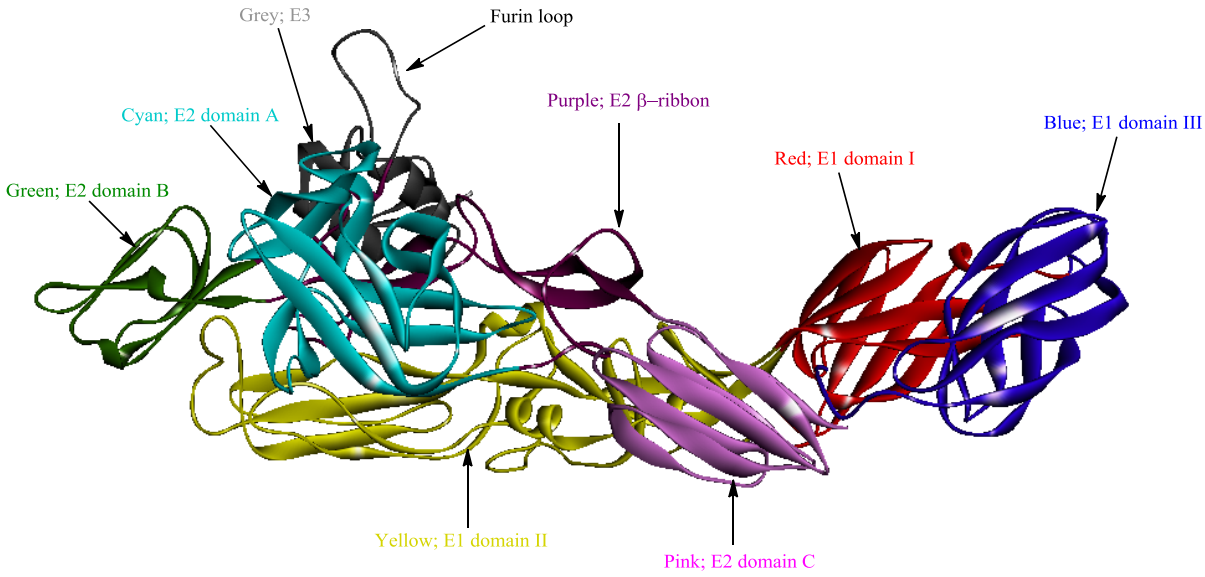


Figure 1.10 Crystal structure of the immature envelope glycoprotein complex of Chikungunya virus, E1 formed of domains I, II and III. E2 contains domains A, B and C, generated from the PDB file code: 3N40.[95] E3 stabilizes the E2 β–ribbon connector[94] being associated with domain A of E2 and domain II of E1, allowing domain B to protect the fusion loop.

The U shaped fusion loop of E1 is inserted in a cavity that lies between E2 domains A and B that is stabilized by hydrogen bonds (Figure 1.13) with E2 histidine side chains.[95] At neutral pH, E3 maintains the relative orientation of the E2 domain B and the domain A, creating the cavity space accommodating the E1 fusion loop. This orientation by E3 protects the virus from premature fusion with other cellular membranes.[95,101] The fusogenic activity of the E1 fusion peptide is therefore, highly dependent on pH change. The histidine residues of E2 act as the pH sensor for the activation of the fusion protein at lower pH[95] due to the increased probability of histidines to become positively charged at lower pH values (acidic endosome), based on

the fact that the imidazole ring of the histidine residue is the only amino acid side chain whose apparent dissociation constant from protons (pKa; acid dissociation constant at logarithmic scale) falls within the physiological range.

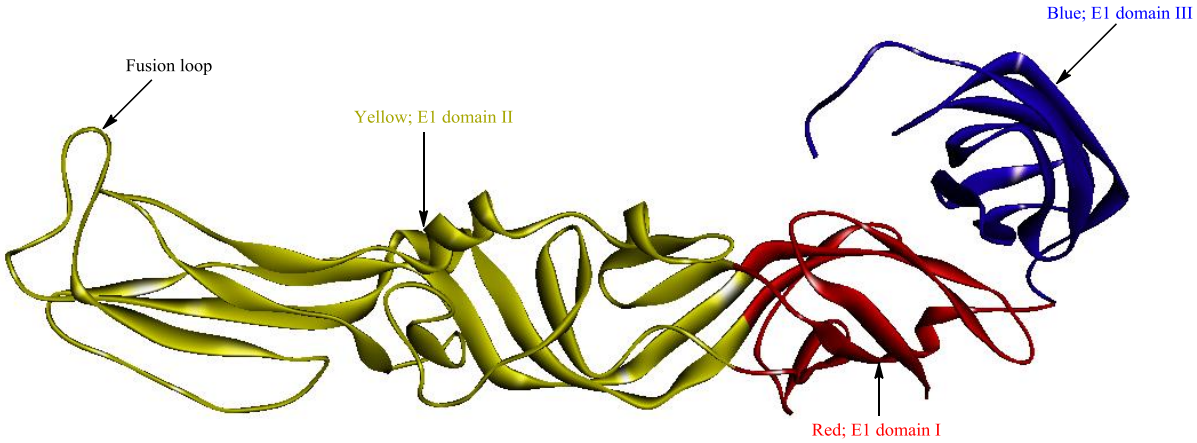


Figure 1.11 Crystal structure of E1 glycoprotein formed of 393 residues (residue 402-residue 794 within the whole complex structure), separated from the complex for visualization, generated from the PDB file code: 3N40.[95]

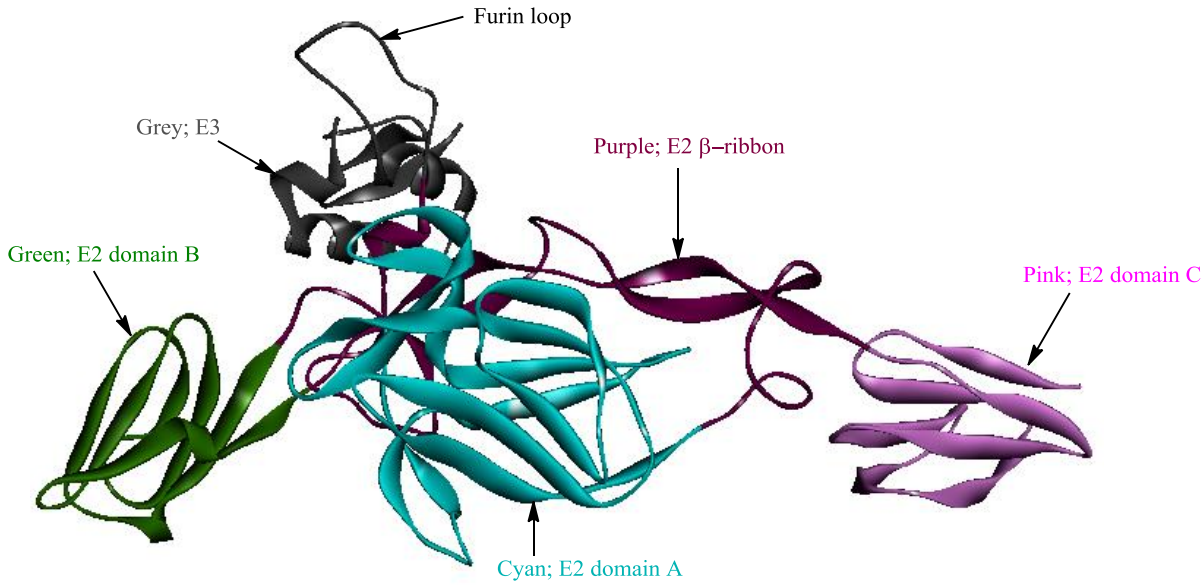


Figure 1.12 Crystal structure of E2, E3 glycoprotein complex formed of 401 residues (residue 1-residue 401 within the whole structural protein complex), separated from the complex for visualization, generated from the PDB file code: 3N40.[95]

Some important locations (transitional epitopes) were identified in both E1 (domain III) and E2 (domain B) of *alphaviruses*: these locations become accessible

upon exposure to heat or low pH.[102,103] These locations are also accessible upon the contact of the virions with the susceptible cells. This contact leads to conformational changes related to cell binding,[104] e.g. domain B moves out in relation to domain A opening the cavity, and the fusion loop now becomes free to release.[95] This occurs without a full dissociation of the E2-E1 heterodimer. Other residues in domain B of E2 are believed to be associated with cell recognition assuming that several sites on the virus surface can interact with different cell receptors and therefore, these residues may be involved in the attachment and entry of the virus.[94]

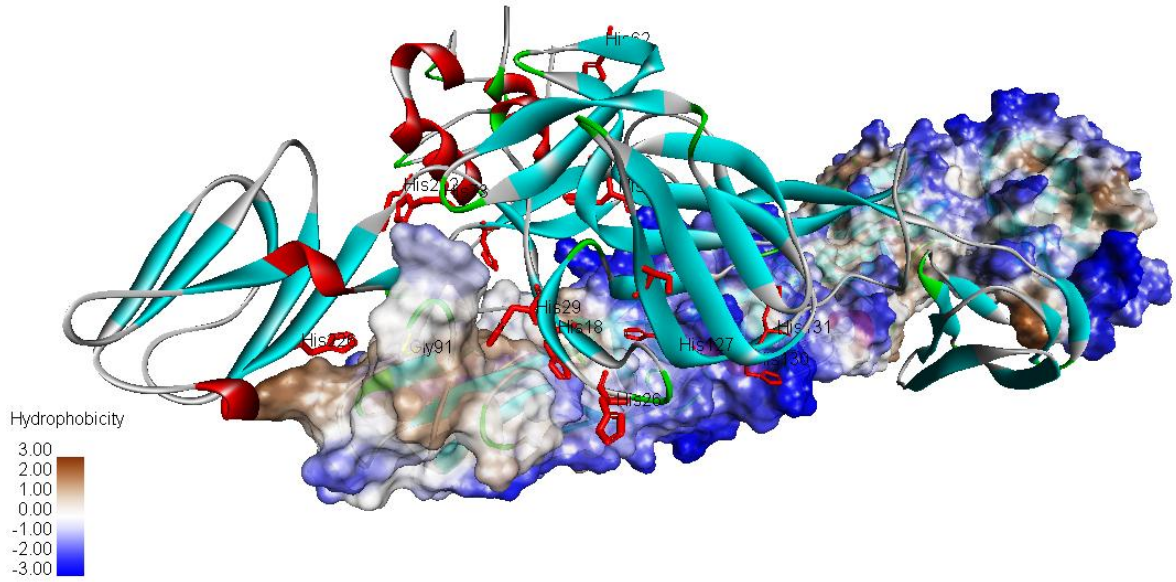


Figure 1.13 Crystal structure of the CHIKV mature envelope glycoprotein complex. E1 is viewed as hydrophobic surface with the critical amino acid residue Gly91[99] of the fusion loop labelled in black. E2 and E3 are viewed as solid ribbon, with histidine residues on E2 viewed as stick structures in red surrounding the fusion loop and acting as the pH sensors[95], generated from the PDB file code: 3N42.[95]

1.5. Highlights for the CHIKV emerging targets

With the recent growing knowledge and available structural information about the CHIKV genome, drug design of particular inhibitors targeting individual viral proteins has become more real. The most promising targets from a chemical and biological standpoint would be the viral protease (nsP2) and the viral envelope proteins. The first protein has a complete crystal structure for the protein N and C domains, while the critical residues for the proteolytic activity have already been investigated. This nsP2

protein also functions through the N domain, and therefore, it would be interesting for medicinal chemists to investigate all possible options to inhibit the function of this viral protein within the two domains. The challenge here is the lack of a tool to validate the inhibitory effects of the designed protein antagonists, rather than whole cell assay protocols. Therefore, extensive efforts should be devoted to investigate the use of the nsP2 as an applicable target for the structure based drug design. The second promising protein target is the viral envelope proteins, with several residues identified to be essential for the viral fusion process such as the Gly91 and His230 residues. Designing specific inhibitors targeting the viral fusion process would be valuable for the inhibition of the *alphaviruses* in general. The viral nsP3 also represents a possible drug design target, with the structure of the conserved macro domain already known. However, this target requires further investigation, such as the complete protein structure as well as a tool to validate the inhibitory activity of the designed antagonists.

1.6. Development of chemotherapeutics against CHIKV: new medicinal chemistry leads

There is currently no recognised single antiviral treatment for chikungunya. During the recent outbreaks that occurred in the Indian Ocean nations, only surface treatments were available, based on non-steroidal anti-inflammatory, non-salicylate analgesics and fluids. Mild physical exercise is believed to decrease the joint stiffness, but heavy exercise may increase the rheumatic pain. During the chronic CHIKV infection, corticosteroids may be used to help decrease the inflammation.[105] The status of drug discovery for the CHIKV is still in the early stages with no drugs currently in clinical trials. The first mouse model to study the pathophysiology of the resulting disease was developed in 2008,[106] after which several animal models were developed to aid the understanding of the drug-disease interactions that would facilitate the development of an effective therapy.[107]

1.6.1. Protease inhibitors

Targeting the CHIKV nsP2 protease activity within the C-domain, would have an inhibitory effect on viral replication. Using the developed homology model for the nsP2, Singh and coworkers screened, *in silico*, a library of compounds to identify potential hits for the predicted active site. Four compounds **(1-4)** shown in Figure 1.14, were identified as potential *in silico* inhibitors for the nsP2 protease.[69] Ideally, binding

to this active site will block the protein function and will stop the replication cycle. However, the actual antiviral activity of these hits **(1-4)** has not yet been investigated.

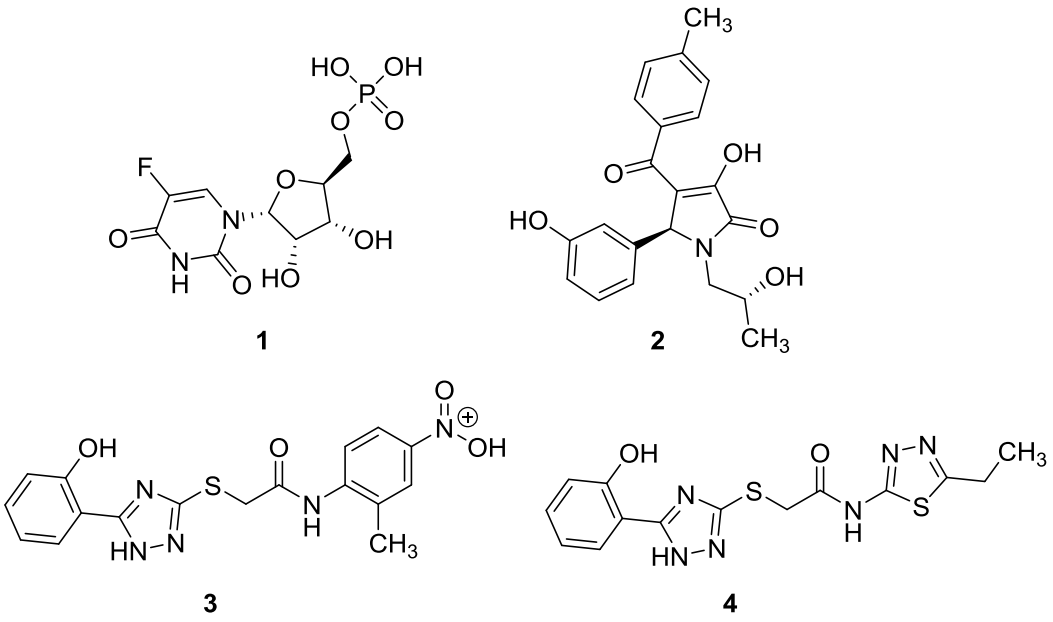
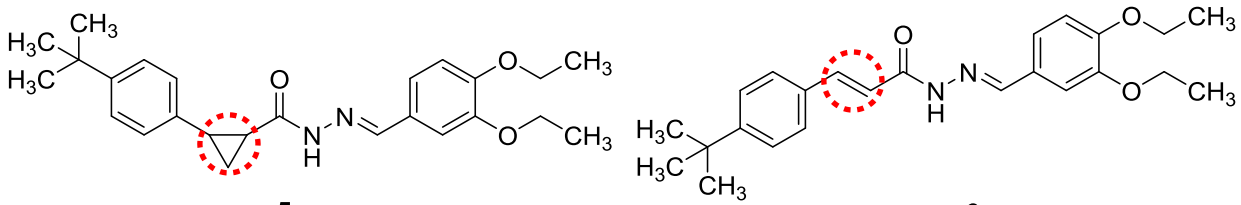


Figure 1.14 Chemical structure of the four identified *in silico* CHIKV nsP2 inhibitors.

In similar work, Bassetto *et al.*[71] reported the identification of *in silico* CHIKV nsP2 inhibitors through virtual screening of a large compound library using the developed homology model for the CHIKV nsP2. One of the hits, compound **5** (Figure 1.15) was predicted to bind to the central portion of the nsP2 protease active site, with its hydrazone group placed in the region defined by the catalytic dyad. The *in vitro* activity of this hit compound was assessed and it was found to inhibit the virus at EC_{50} value of 5 µM and selectivity index (SI) value of 14, through the inhibition of the virus-induced cytopathic effect.[71] The central cyclopropyl and the hydrazone moieties were found, through a structure activity relationship study, to be important for the anti-chikungunya activity. A series of derivatives were also designed based on these identified pharmacophores of **(5)**. When the cyclopropyl group was replaced with a trans-ethenylic moiety maintaining length and geometry of the original linker (**6**, Figure 1.15), the antiviral activity was slightly improved. Compound **(6)** displayed EC_{50} value of 3.2 µM and selectivity index (SI) of 32. The binding modes of both compounds (**5**, **6**) inside the nsP2 pocket were similar.[71]



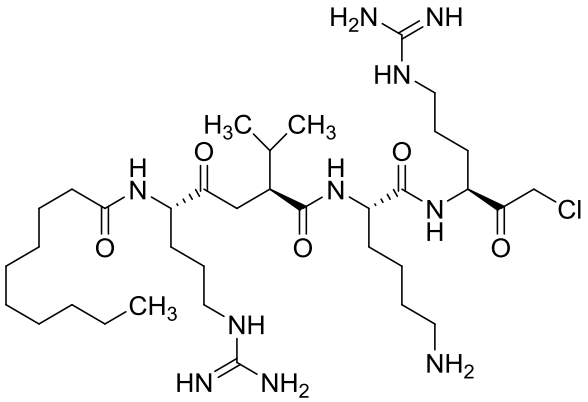
5 **6**

EC_{50}= 5 μM EC_{50}= 3.2 μM

Figure 1.15 Chemical structure of the CHIKV inhibitors identified from the *in silico* screening, red circles indicates the only difference in the structures responsible for the activity change.

1.6.2. Furin inhibitors

Infection by *alphaviruses* can be inhibited *in vitro* by blocking the intracellular furin-mediated cleavage of viral envelope glycoproteins: the E2E3 or p62 precursors. This blocking was demonstrated by showing the inhibitory effect of an irreversible furin-inhibiting peptide, decanoyl-RVKR-chloromethyl ketone (dec-RVKR-cmk, **7**, Figure 1.16) on *in vitro* CHIKV infection.[100] This peptide significantly reduces the processing of E3E2 CHIKV glycoproteins in infected myoblast cultures and led to the formation of immature viral particles and impaired viral spreading among cells, but not the replication in the already infected cells.[100] Therefore, the chemical structure of the furin-inhibiting peptide **(7)** could be a starting point for generating novel generations of active peptidomimetics using the ligand based drug design techniques, targeting the intracellular furin cleavage step.

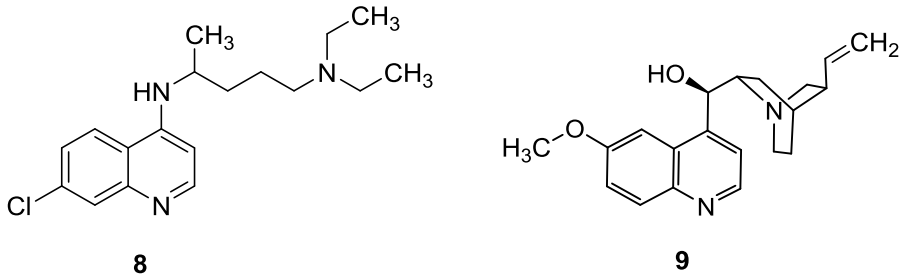


7

Figure 1.16 Chemical structure of the furin-inhibiting peptide **(7)**.

1.6.3. Chloroquine and Quinine

The *in vitro* antiviral activity of chloroquine (**8**, Figure 1.17) was first reported more than 35 years ago and has been successfully used as anti-malarial drug.[108,109] Concerning the *alphaviruses*, chloroquine was found to be effective *in vitro*,[110,111,112] but a few years later, a mouse model revealed that chloroquine may enhance viral replication *in vivo* leading to aggravation of the disease.[113] Regarding the CHIKF, chloroquine and chloroquine phosphate have been used in the treatment of chronic chikungunya arthritis,[114] only for the anti-inflammatory properties of the molecule (used in chronic rheumatologic diseases) rather than for the antiviral effect. Some studies suggest that chloroquine might interact with the endosome-mediated internalisation process during the infection cycle, stating that chloroquine might be classified as an entry inhibitor. Chloroquine **8** entered phase 3 clinical trials in France as a therapy for the CHIKV in 2006, however, these studies were terminated in 2007 without promising results.



8 **9**

IC_{50}= 1.1 μg/ml IC_{50}= 0.1 μg/ml

Figure 1.17 Chemical structures of chloroquine (**8**) and quinine (**9**).

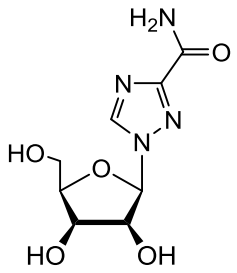
In one clinical study, the effect of chloroquine **8** on CHIKF patients was investigated using another placebo treated group of patients, and at the conclusion of the study period, there was no statistical difference between the chloroquine and the placebo treated groups regarding the mean duration of febrile arthralgia or the decrease of viraemia (viral count in plasma). However, these contradicting results casts serious doubt as to the effectiveness of chloroquine as an effective chemotherapeutic against CHIKV and until resolved, it should be treated with caution as a drug lead.[105]

Another anti-malarial drug, quinine (**9**, Figure 1.17), also inhibits the virus *in vitro* in a concentration less than that of chloroquine (IC_{50} = 0.1 μg/ml for quinine, 1.1

µg/ml for chloroquine). Also, quinine is suggested to affect the nsP1 as mutations in this protein occur upon growing the virus in a high concentrations of quinine.[105]

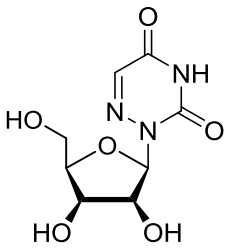
1.6.4. Ribavirin and 6-Azauridine

The antiviral ribavirin (**10**, Figure 1.18), is well known to inhibit many RNA viruses *in vitro*, by different mechanisms.[115] It is being used either alone for treatment of Respiratory Syncytial virus, or in combination with alpha-interferon in the treatment of hepatitis C virus (HCV) infection. This combination showed a synergistic *in vitro* inhibition of the CHIKV.[116] A combination of alpha-interferon, at a concentration of 3.9 IU/ml, and ribavirin at a concentration of 18.75 µg/ml, inhibited CHIKV replication by 50%, whereas ribavirin **10** alone without interferon, inhibited CHIKV with EC_{50} value of 83.3 µg/ml. However, there is no evidence supporting the clinical efficacy of ribavirin **10** on CHIKV, and the combination with interferon should be subjected to clinical trials for the treatment of CHIKV infections.[105] Furthermore, the exact mechanism of ribavirin is still unclear as it may change from one virus to another, however, it is believed that ribavirin can interact with the intracellular viral RNA production.



10

EC_{50}= 83.3 µg/ml



11

EC_{50}= 0.2 µg/ml

Figure 1.18 Chemical structures of ribavirin (**10**) and 6-azauridine (**11**).

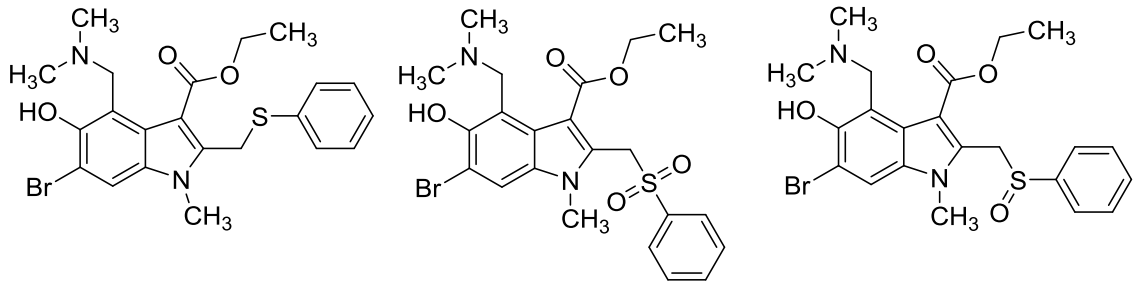
The broad-spectrum anti-metabolite, 6-azauridine (**11**, Figure 1.18) inhibits both DNA and RNA virus replication, and the activity might be through the inhibition of orotidine monophosphate decarboxylase, an enzyme involved in the *de novo* biosynthesis of pyrimidine, cytidine and thymidine.[117] It showed a significant inhibition of CHIKV at a low concentration (0.2 µg/ml) and was more effective against the CHIKV compared to ribavirin **10**.[116] 6-Azauridine **11** is not approved for human use and therefore was not included in a combination study with alpha-interferon. However,

6-azauridine triacetate is used for treatment of different diseases without notable adverse effects.[118] Therefore, 6-azauridine should be evaluated *in vivo* as CHIKV inhibitor.[116]

1.6.5. Arbidol

The antiviral drug arbidol (**12**, Figure 1.19) was originally developed 20 years ago at the Russian Research Chemical and Pharmaceutical Institute.[119] Since 1990, it has been used in Russia for acute respiratory infections including influenza. So far, arbidol shows a wide range of activity against many RNA, DNA, enveloped and non-enveloped viruses.[120] This broad spectrum of activity may be attributed to the different modes of action including the inhibition of virus mediated fusion,[121] and blocking of the viral entry into the target cells through inhibition of glycoprotein conformational changes that are essential for the fusion process, as in the case of influenza virus and hepatitis C virus.[122,123]

In 2011, arbidol **12** and two derived metabolites (Figure 1.19), HZ1 **13** and HZ3 **14**, were tested *in vitro* on the chikungunya virus using two cell lines, and in different conditions, pre- and post-infection treatments. Arbidol was the only active compound with an IC_{50} value that was much lower than the toxic concentration (IC_{50} = 12.2 µM, $CC_{50} \geq 200$ µg/ml).[124] HZ1 **13** and HZ3 **14** were assumed to be responsible for the anti-viral properties of arbidol, however in this study, HZ1 and HZ3 showed only weak effects on CHIKV replication.



12 **13** **14**

IC_{50}= 12.2 µM IC_{50}= 56.4 µM IC_{50}= 54.5 µM

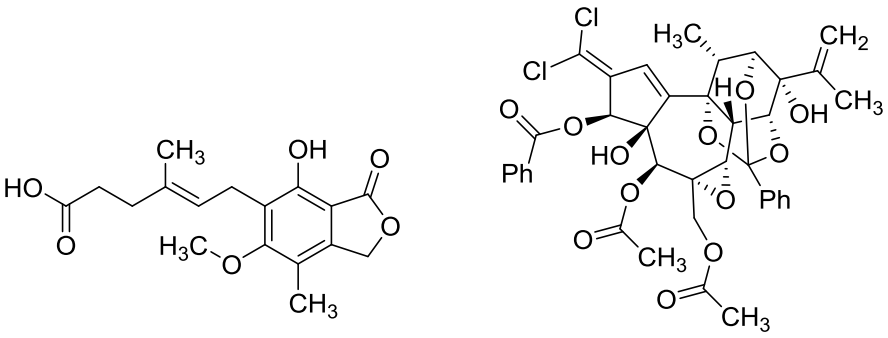
Figure 1.19 Chemical structures of arbidol (**12**), HZ1 (**13**) and HZ3 (**14**).

Arbidol **12** was found not to be viricidal, rather, it blocks the earliest stages of the viral replication, virus attachment and/or virus entry as previously reported.[120] For

CHIKV, it works by targeting the cellular membrane (E2 viral envelope protein) which was confirmed by the use of arbidol resistant CHIKV strain (mutation in the E2 protein, G407R).[124] However, all these studies were performed *in vitro* and *in vivo* studies are required to validate the activity of arbidol on CHIKV.

1.6.6. Mycophenolic acid (MPA)

Mycophenolic acid (**15,** Figure 1.20) was isolated about one hundred years ago.[125] It is an inhibitor of inosine monophosphate dehydrogenase (IMPDH), an enzyme involved in the *de novo* biosynthesis of guanine nucleotide. It has a very good anti-proliferation activity and has been established as an anticancer,[126] antiviral agent and as an immunosuppressant.[127]



15 **16**

IC_{50}= 0.2 μM EC_{50}= 1.5 μM

Figure 1.20 Chemical structures of mycophenolic acid (**15**) and trigocherrin A (**16**).

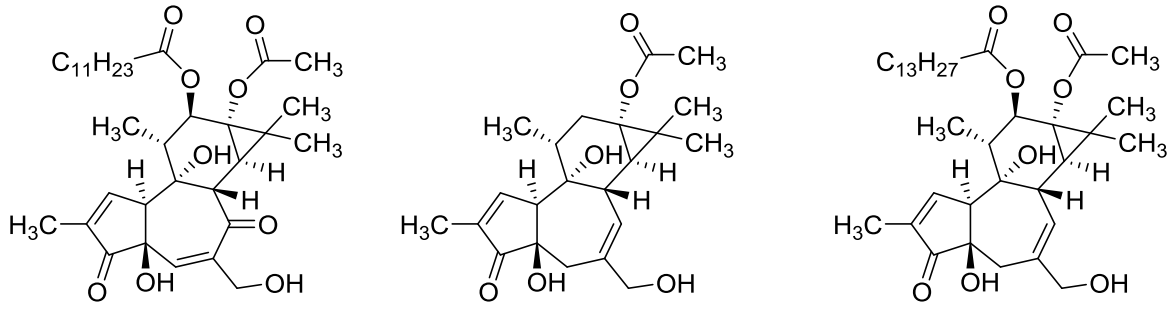
Recently, MPA was shown to inhibit the CHIKV replication and virus induced cell death. The IC_{50} value was 0.2 μM with a selectivity index of 150.[128] Mycophenolic acid was found to induce CHIKV apoptosis. When the treated CHIKV infected cells were provided with exogenous guanosine (GMP), MPA could no more prevent the CHIKV induced cell death, indicating that it works by inhibiting IMPDH enzyme. It has also been suggested that MPA increases the mutation rate during the viral replication.[128] Considered as a good lead compound, the *in vivo* activity of mycophenolic acid on CHIKV should be investigated as it is well known that mycophenolic acid suffers from a metabolic drawback associated with rapid conjugation of the C-7 phenolic hydroxyl group with glucuronic acid.[129,130]

1.6.7. Trigocherrin A

Trigocherrin A (**16**, Figure 1.20), is a highly oxygenated and chlorinated daphnane diterpenoid orthoester, that had been isolated from the bark of *Trigonostemon cherrieri* Veillon (Euphorbiaceae), a tree collected in the sclerophyllous forest of New-Caledonia. This genus comprises about 80 species occurring in tropical Asia, from India and Sri Lanka to New Guinea.[131] These diterpenoids have been shown to have cytotoxic[132,133] and antiviral[134,135] properties. Recently, this natural product was tested against CHIKV and was found to inhibit virus-induced cell death in a virus-cell-based assay with an EC_{50} of 1.5 μM and only caused a significant antimetabolic effects at a concentration (CC_{50}) of 35 μM, and the selectivity index (SI) value was 24. Different concentrations of **16** were able to protect the host cells from the virus cytopathic effect without any adverse effects. Also, it was found to be more potent than the reference compound used in this study (chloroquine, **8**).[136,137] The assay protocol suggests that this compound might inhibit the viral replicase functions, however, for the drug discovery process, the *in vivo* activity, the precise mode of action as well as the total synthesis of this lead compound should be investigated.

1.6.8. Trigowiin A, Prostratin and 12-O-Tetradecanoylphorbol 13-Acetate

In a recent study (2012), an extract of the bark of *Trigonostemon howii* (Euphorbiaceae species) from central Vietnam was tested against the CHIKV.[138] A new tigliane diterpenoid, trigowiin A (**17**, Figure 1.21) was isolated and was found to be structurally closely related to the tigliane diterpenes (Figure 1.21**)**, prostratin (a promising adjuvant for anti-HIV therapy)[139] **18** and 12-*O*-tetradecanoylphorbol 13-acetate (TPA, **19**).[138]



17 **18** **19**

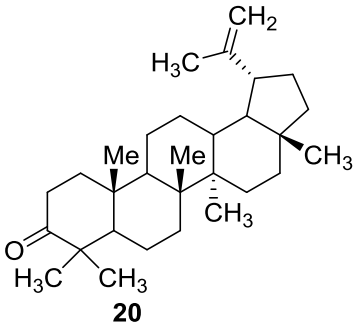
EC_{50}= 43.5 μM EC_{50}= 2.6 μM EC_{50}= 0.0029 μM

Figure 1.21 Chemical structures of trigowiin A (**17**), prostratin (**18**) and TPA (**19**).

In the CHIKV assay, trigowiin A **17** showed weak antiviral activity, with an EC_{50} of 43.5 ± 12.8 μM whereas TPA **19** and prostratin **18** proved to be the most potent inhibitors, with low EC_{50} values and higher selectivity indices (SI), for TPA **19**, EC_{50} = 0.0029 ± 0.0003 μM and SI = 1965 while EC_{50} for prostratin **18** was 2.6 ± 1.5 μM and SI = 30.3 showing that TPA **19** was 65 times more potent than prostratin **18.**[138] At the same time, these compounds did not show activity against Sindbis virus (SINV) nor Semliki forest virus (SFV) which indicates an excellent selectivity for inhibition of the CHIKV. The authors[138] believe that this selectivity might be due to a specific mechanism of virus inhibition through the activation of the signal transduction enzyme protein kinase C (PKC), similar to the proposed mechanism of HIV replication inhibition[140] for TPA **19**. However, TPA **19**, is currently one of the most potent tumor-promoting agents known to date,[141] therefore, the less potent prostratin **18** is more likely to be further investigated as a potential candidate, as it does not exhibit[142] a tumor-promoting activity. The exact mode of action of these compounds is as yet unknown and needs further investigation.

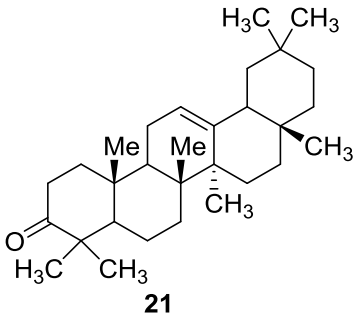
1.6.9. Lupenone and β-amyrone

In the continuous effort to identify novel inhibitors of Chikungunya from natural sources, a phytochemical study on the leaves of *Anacolosa pervilleana* (Madagascan plant) was performed in a virus-cell-based assay for CHIKV.[143] Two triterpenoids with a moderate anti-CHIKV activity were isolated and identified (Figure 1.22), lupenone **20** with EC_{50} = 77 μM and β-amyrone **21** with EC_{50} value of 86 μM.[143] However, the weak activity of these natural compounds should be improved by medicinal chemistry optimization processes before being considered as lead compounds.



20

EC_{50}= 77 μM



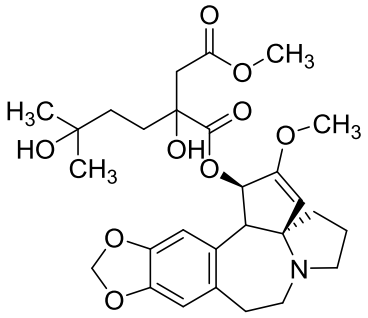
21

EC_{50}= 86 μM

Figure 1.22 Chemical structures of lupenone (**20**) and β-amyrone (**21**), Me refers to methyl groups, simplified to fit within the ring structure.

1.6.10. Harringtonine

Harringtonine (**22**, Figure 1.23) is an alkaloid from *Cephalotaxus harringtonia* trees, which is native to Japan. The natural product has known antitumor activity and inhibits the first cycle of the elongation phase of eukaryotic translation.[144] Recently, the anti-CHIKV activity of harringtonine was investigated.[145] It displayed potent inhibition of CHIKV infection (EC_{50} = 0.24 μM) with minimal cytotoxicity, through the inhibition of the early stages of infection after cellular endocytosis. Also, it was found to affect the CHIKV RNA production inside the infected cell, as well as viral protein expression such as the nsP3 and the E2 proteins.[145] The *in vivo* studies of harringtonine are still ongoing which could make harringtonine a promising lead towards the discovery of anti-CHIKV drugs.



22

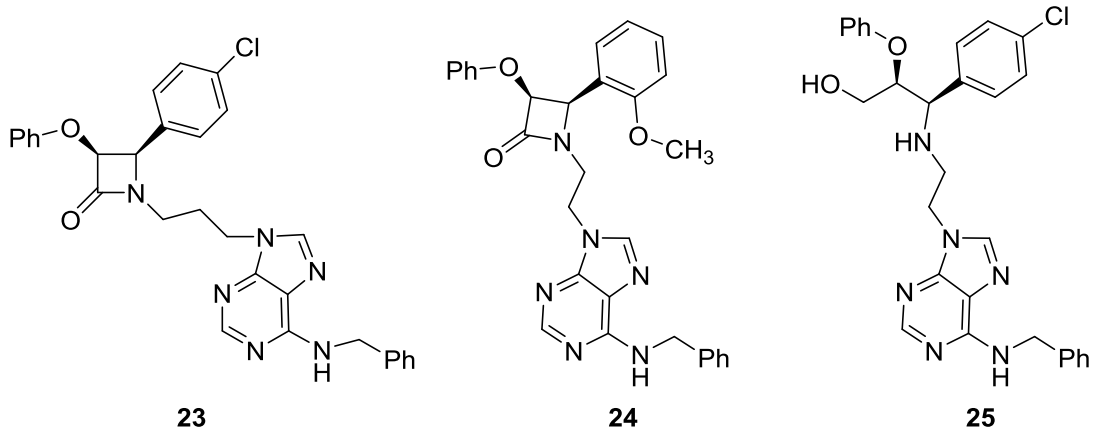
EC_{50}= 0.24 μM

Figure 1.23 Chemical structures of harringtonine (**22**).

1.6.11. Purine based inhibitors

In 2012, D'hooghe *et al*. reported the design and synthesis of a new series of purine-*β*-lactam hybrids and purine-aminopropanol hybrids and their evaluation as potential antiviral candidates depending on the antiviral templates purines and *β*-lactams.[146] These new scaffolds were screened against nine different viruses including the chikungunya virus. Two purine-*β*-lactam hybrids and one purine-aminopropanol hybrid (Figure 1.24) were found to possess promising activity and cytotoxicity profiles, the purine-*β*-lactam **23** with EC_{50} = 17.11 μM and SI > 5.75, the purine-*β*-lactam **24** with EC_{50} = 13.01 μM and SI > 4 and the purine-aminopropanol **25** with EC_{50} = 11.51 μM and SI > 6, which displayed relatively stronger inhibition compared to the β-lactam ring

analogues, indicating that the β-lactam ring is not essential.[146] The mode of action has not been investigated. The synthesis of this class of compounds is already established and therefore, they represent good subject for further medicinal chemistry optimization.



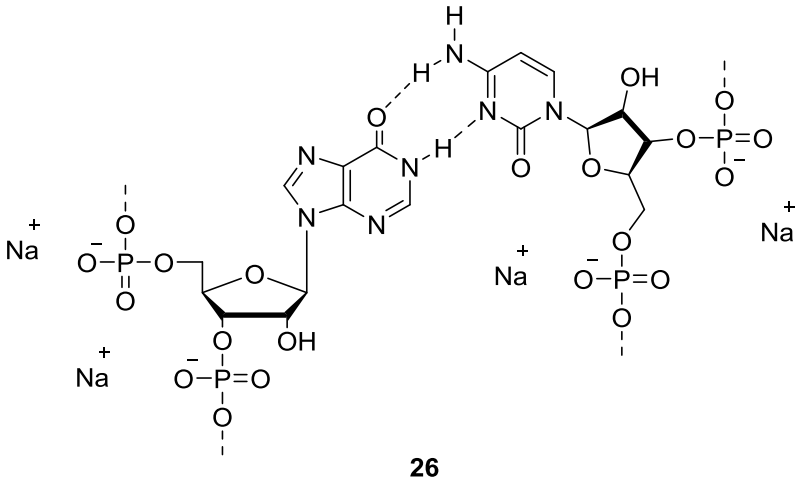
23 **24** **25**

EC_{50}= 17.11 μM EC_{50}= 13.01 μM EC_{50}= 11.51 μM

Figure 1.24 Chemical structures of the purine-*β*-lactams (**23**, **24**) and the purine-aminopropanol (**25**) compounds.

1.6.12. Polyinosinic acid

Polycytidylic acid [Poly (I:C)] (**26**, Figure 1.25), a synthetic double-stranded RNA (dsRNA) analogue, is an immunostimulant that acts as an inducer for the most potent interferon (IFN) *via* interaction with toll-like receptor 3 (TLR3).[147] It can induce IFN-α/β production and natural killer (NK) cell activation *in vivo* after intraperitoneal injection.[148] Activation of the TLR3 contributes to an innate immune response against many viruses.[149] In CHIKV infection, the virus was found to be sensitive to the innate immune response induced by Poly (I:C). This sensitivity was noticed as a decreased cytopathic effect and inhibition of the virus replication in the infected cell lines. This sensitivity has been explained to be a result of the overstimulation of the TLR3 as well as the other anti-viral genes by Poly (I:C).[150]



26

Figure 1.25 Chemical structures of polycytidylic acid (**26**).

1.6.13. Gene silencers

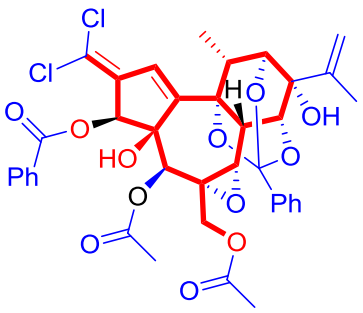
New trends in the CHIKV treatment trials are the use of genes silencers targeting specific viral proteins (capsid protein, E1, nsP1 and nsP3). Silencing the target viral genes will consequently lead to shutdown of the protein expression process, and thereby stopping viral replication. After viral infection, the exogenous small interfering RNA (siRNA) induces RNA interference mechanism, resulting in the assembly of RNA-induced silencing complex (RISC) which inhibits subsequent protein expression. Using the siRNA targeting CHIKV E1 and nsP3, was effectively used in suppression of *in vitro* CHIKV replication.[151] Similarly, a plasmid based small hairpin RNA (shRNA) against CHIKV replication targeting the capsid, E1 and nsP1 proteins has been used. Simply, the plasmid is first introduced to cell, expressed inside the nucleus, small-hairpin RNA (shRNA) is formed and processed by cytoplasmic Dicer enzyme to siRNAs, leading to activation of the RNA silencing machinery. This silencing machinery recognizes and degrades the target CHIKV single strand RNA, consequently stopping viral protein expression.[152] However, clinical studies should be able to prove the applicability of these trends in developing effective anti-CHIKV therapeutics.

1.7. Highlights for CHIKV inhibitors

Several molecules have been tested against the emerging CHIKV with weak to moderate activities. Those included drugs already in the market, being used for other diseases, such as chloroquine **8**, ribavirin **10**, arbidol **12** and mycophenolic acid **15**. The

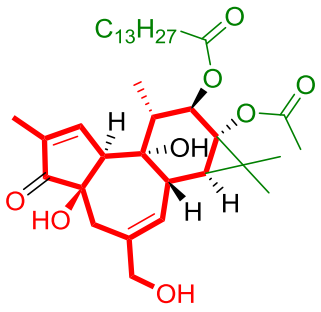
challenge here would be in finding similar compounds that are structurally unique with improved potency and other drug like qualities, also avoiding the metabolic draw backs when present. The protease inhibitors **5**, **6** that were developed based on an *in silico* study, would be attractive candidates for further development, however, the challenge here is to develop a more specific inhibitory assay protocol to confirm the selectivity for the inhibition of viral proteins. Protein crystallization with the inhibitor would be a useful element to validate such studies. Some of the tested compounds represent complex natural products and have good enough activity as antiviral agents, such as Trigocherrin A **16**, prostratin **18**, TPA **19** and harringtonine **22**. It will be an arduous task to chemically access these structures and to simplify these chemical skeletons to more drug like molecules with acceptable ADME properties.

The anti-CHIKV activity of the tested molecules ranged from strong to weak inhibition depending on the type of the assay used, with the TPA **19** being the strongest inhibitor with EC_{50} = 0.0029 μM and β-amyrone **20** which displayed the weakest activity with EC_{50} = 86 μM, similar activity to that of ribavirin **10**, EC_{50} = 83.3 μM. It is worth noticing that the active agents, TPA **19** and the less active Trigocherrin A **16**, shared a common structural feature, the substituted benzo[*e*]azulene derived structure, (Figure 1.26). The TPA structure **19** is simpler than that of Trigocherrin A **16**, lacking the extra phenyl rings, two chlorine atoms, four oxygen bridges and the alkene side chain (Figure 1.26), however, TPA **19** on the other hand has a characteristic long tetradecanoic ester moiety which was responsible for the activity over Trigowiin A **17** and Prostratin **18**, two derivatives that were even less active than Trigocherrin A **16**. These highlights will guide further investigations toward the development of CHIKV inhibitors based on that substituted benzo[*e*]azulene skeleton.



16

EC_{50}= 1.5 μM



19

EC_{50}= 0.0029 μM

Figure 1.26 Chemical structures of TPA (**19**) and Trigocherrin A (**16**), common skeleton is highlighted in bold red colour, unique groups to TPA (**19**) are shown in green while the excess groups of Trigocherrin A (**16**) are shown in blue colour.

1.8. Concluding remarks

Chikungunya virus (CHIKV) is an emerging arbovirus virus that has had devastating effects in recent years in many areas in the world. Chikungunya virus infection can develop into an arthritis disease that remains with the patient for years. The mutations in the viral envelope protein genes increased the fitness of the virus in another mosquito vector, *Ae. albopictus* which was responsible for cases reported in temperate zones. With no licensed vaccine for immunization against this virus, disease control is currently non-existent, and the treatment would be through the development of chemotherapeutics. Some promising lead compounds have been discovered recently and could be starting points towards effective treatments. The discovery of these leads was mostly based on random screening of drugs already on the market, newly discovered natural products or the antiviral evaluation of synthetic compounds. Also, in the last two years, the molecular functions as well as the crystal structures of a number of critical enzymes involved in the virus life cycle are being reported, with no known inhibitor thus far. Some *in silico* models targeting these viral proteins have been developed, with some active compounds identified, to initiate the search for effective and selective therapy such as targeting the viral protease. This will help the drug discovery and development process through designing inhibitors against those targets. The virus has been considered neglected for many years and this emphasizes the importance of developing highly potent and safe inhibitors.

1.9. Project aims

At the outset of this project (2010), there were no lead compounds reported (including those previously mentioned in this chapter **1**-**26**), except for some contradicting arguments about the efficacy of chloroquine **8**, with limited evidence suggesting the efficacy against the chikungunya virus. There were no crystal structures available for any of the viral proteins. Therefore, the main aim of the project was to find a hit compound that could be taken further *via* a hit to lead discovery study. The main strategies investigated during this project were:

- Random screening, through the synthesis and evaluation of anti-microbial class of compounds that were driven from the current projects within our research group.
- Continuous screening of the literature for emerging viral proteins to develop a structure-based drug design approach.
- Continuous screening of the literature for active lead compounds to conduct a ligand-based drug design approach.

2.1. Introduction

At the outset of this project (2010), there were no lead compounds reported, except some contradicting arguments about the efficacy of chloroquine **8** against the chikungunya virus, as was referred to in chapter 1. Therefore, the goal of this project was to search for new lead compounds. Finding a lead compound usually starts with the random screening of a compound library against the target virus. One active program within our group was the investigation of new structural motifs as potential non-nucleoside inhibitors of HIV-1 reverse transcriptase (HIV RT)[153] and developing *in silico* models for the identification of structurally unique inhibitors. While 15 compounds tested from this *in silico* screening of these models were inactive against HIV RT,[154] seven possessed good anti-malarial activity. Of these seven compounds, the arylarenepyrimidylmethane (AAPM) compound **27** (Figure 2.1) was chosen for further investigation as an anti-malarial lead, where it showed an IC_{50} of 0.88 µM against *P. falciparum*.[155,156] The developed synthetic strategy utilized for the synthesis of compound **27** is shown in Figure 2.1.

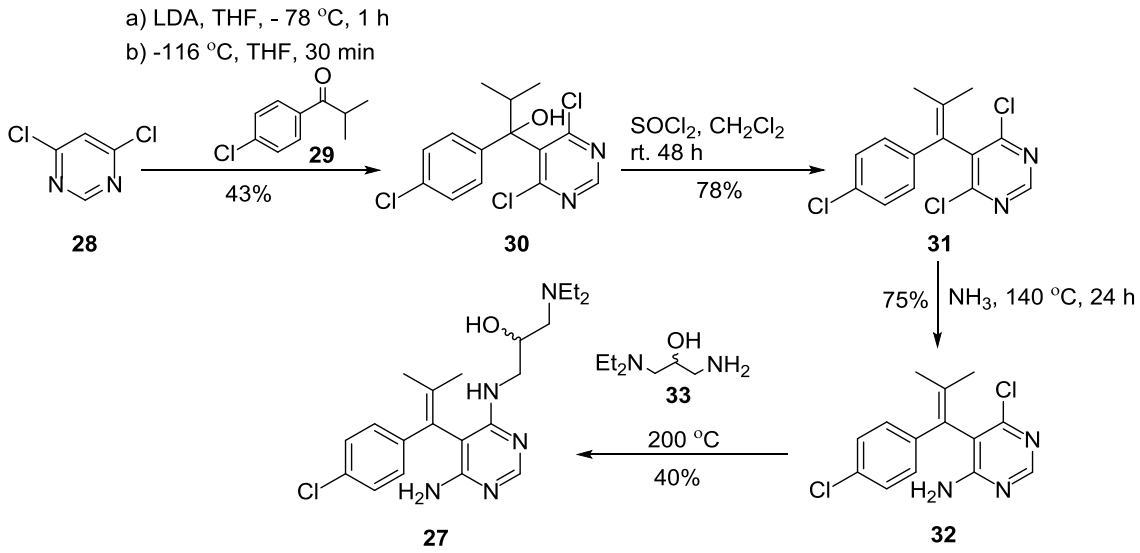


Figure 2.1 The synthesis of the AAPM derivative, compound **27**.

When compound **27** was synthesized, no similar derivatives were investigated to study structure activity relationships for the anti-malarial activity, nor an optimization strategy for the synthetic pathway used to obtain more derivatives. Therefore, our research group had focused on developing this series, not only to further develop the synthetic methodology, but also accessing analogues to investigate the structure activity relationship.

An attempt to derivatise compound **27** in our laboratory had been performed, and only one derivative, the unsubstituted AAPM compound **34** (Figure 2.2) was synthesized, due to the difficulties in the synthesis, even after several trials of optimization.[157]

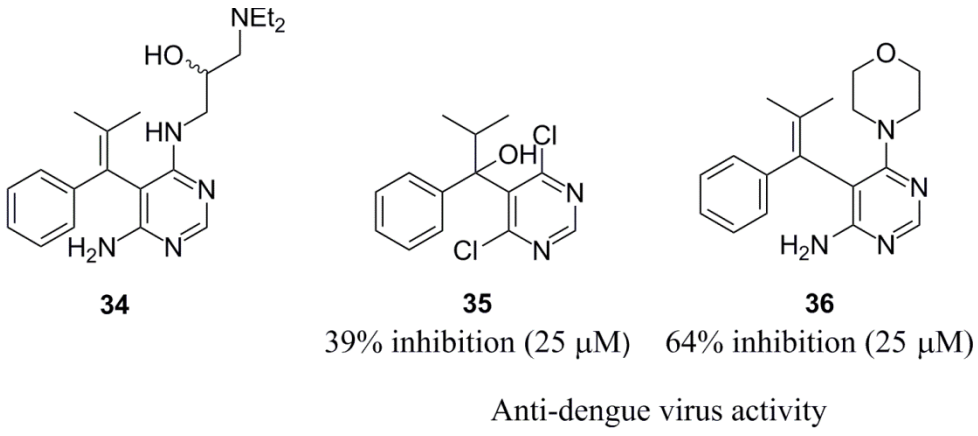


Figure 2.2 The unsubstituted AAPM compound **34**, and the two AAPM modified derivatives **35**, **36** displaying inhibitory activity against the dengue virus close to 50% at 25 μM.

During the optimization[157] process, some modified AAPM derivatives (without the diethylamino-propanol side chain) were synthesized and were tested against dengue virus.[157] The AAPM derivatives **35**, **36** showed weak to moderate activity against the dengue virus (Figure 2.2). Dengue virus belongs to the family Flaviviridae, and is transmitted through the *Aedes aegypti* mosquito,[158] the same vector which transmits the chikungunya virus. Therefore, this family of heterocycles may also inhibit the CHIKV. With an established synthetic strategy, the generation of analogues for testing against the CHIKV makes a viable line of research seeking new leads. Consequently, the synthesized derivatives would enrich the AAPM library giving strength to the scope of

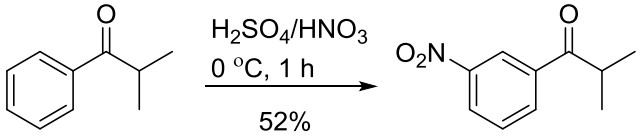
developing novel anti-microbial agents in general, and especially, with respect to the anti-malarial program.

2.2. Synthesis of the *meta*-nitro AAPM derivative

The synthetic strategy for this AAPM derivative was similar to the procedures described in Figure 2.1 replacing the ketone **29** with **38**.

2.2.1. Nitration of isobutyrophenone 37

The synthesis started with the nitration of isopropyl phenyl ketone **37** under the standard nitration conditions of dropwise addition of the cold nitrating mixture to a cold solution of the ketone **37** in conc. sulphuric acid (Scheme 2.1). The isopropyl 3-nitro-phenyl ketone **38** was isolated in 52% yield, as a yellow oil.



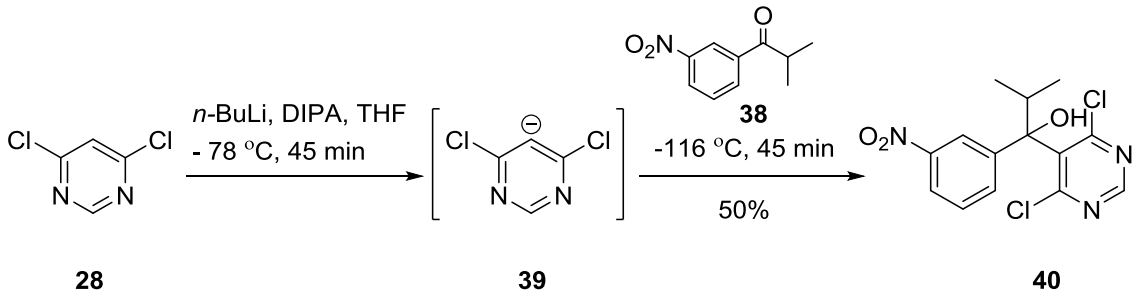
37 **38**

Scheme 2.1 The synthesis of the isopropyl 3-nitrophenyl ketone **38**.

2.2.2. *Addition of the isobutyrophenone 35 to 4,6-dichloropyrimidine 28*

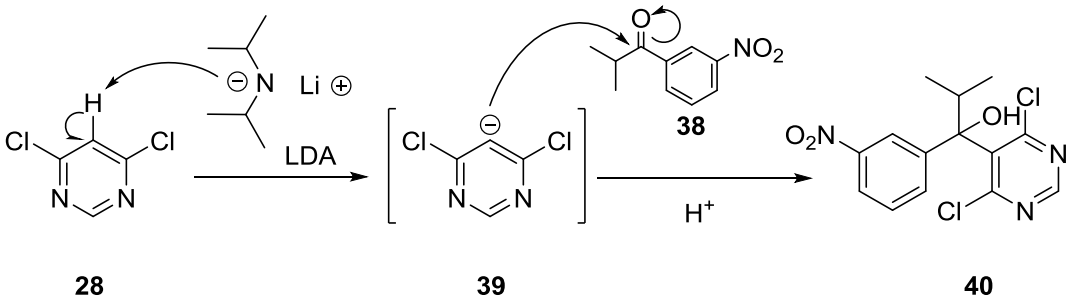
The next step in the synthetic pathway was the addition of the 4,6-dichloropyrimidine **28** to the nitro derivative **38**. Therefore, a fresh solution of lithium diisopropylamide (LDA) solution was prepared by the dropwise addition of *n*-BuLi solution in hexanes to a solution of diisopropyl amine (DIPA) in tetrahydrofuran (THF), pre-cooled to -78 ºC (Scheme 2.2), and the solution was allowed to stir at -78 ºC for 45 min to allow the complete formation of LDA. A solution of 4,6-dichloropyrimidine **28** in THF, pre-cooled to -78 ºC, was then added dropwise to the LDA solution, producing an instant orange colour change which was attributed to the formation of the pyrimidyl anion **39**, and the solution was stirred at -78 ºC for 45 min. The solution was then cooled to -116 ºC, and a solution of nitro derivative **38** in THF, pre-cooled to -116 ºC, was added dropwise. The mixture was left for 45 min to allow complete addition. The reaction was then quenched by addition of NH_4Cl. After acidic aqueous workup, the desired alcohol **40** was isolated in 50% yield as pale yellow crystals, the unreacted starting ketone **38** was recovered with the remaining outcomes of this reaction being polymeric products as analysed from the baseline material on thin layer chromatography (TLC) analysis.

The analysis of the ^{1}H NMR spectrum of **40** showed a broad singlet peak at 3.38 ppm, asssigned to the OH group, and two doublet peaks resonating at 1.11 and 0.93 ppm, that were assigned to the two methyl groups.



Scheme 2.2 Addition of 4,6-dichloropyrimidine **28** to the nitro derivative **38**, forming the alcohol **40**.

The proposed mechanism for this reaction starts with the nucleophilic attack of LDA onto the 5-pyrimidyl proton resulting in the lithiated dichloropyrimidine anion **39**, which nucleophilically attacks the electrophilic carbonyl of the ketone **38**, to give the alcohol after a protic workup (Scheme 2.3).



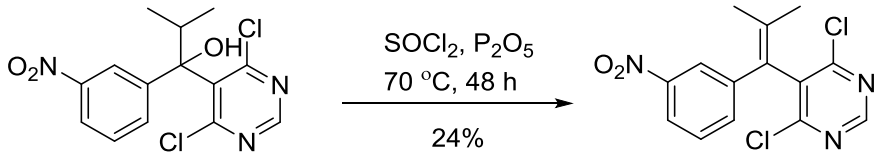
Scheme 2.3 The proposed mechanism for the formation of alcohol **40**.

This reaction (Scheme 2.2) required ultra dry conditions, with all the glassware oven dried for at least 12 h prior to reaction, and all the reactants dried overnight under vacuum. The 4,6-dichloropyrimidine **28** was crystallized from water. Previous work[157] reported the purification of **28** by bulb-to-bulb distillation, however, in this work, simple crystallization was effective. The previous best reported yields[157] of similar reactions in our group was 38-47% yield. In this work (Scheme 2.2), the yield was

optimized to 50%. The notable increase in yield was attributed to the ultra dry conditions, as this reaction was found to be very sensitive to moisture, including from the N_2 atmosphere used during the reaction. Therefore, in this reaction (Scheme 2.2), the needle supplying N_2 to the reaction flask, was filled with pre-dried $CaCl_2$ to ensure the dry atmosphere inside the reaction flask. The slow rate of addition of the electrophile **38**, as well as the pre-cooling (-116 ºC) of the electrophile **38** solution prior to addition, were found to be important factors for improving the yield (Scheme 2.2).

2.2.3. *Dehydration reaction of the alcohol 38 to the form the alkene 41*

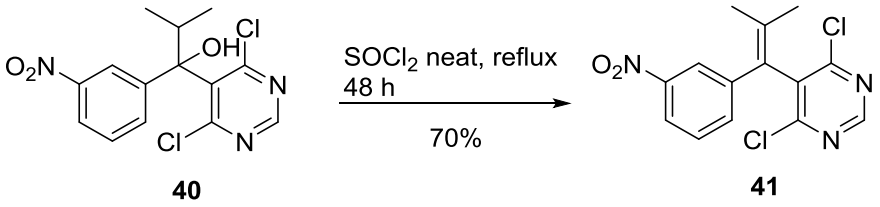
The next synthetic step was the dehydration of alcohol **40**. Previous work[157] facilitaed a similar transformation by reacting the alcohols with four molar equivalents of thionyl chloride ($SOCl_2$) and four molar equivalents of phosphorus pentaoxide (P_2O_5) in dichloromethane with stirring at 50 ºC for 12 h. In this work, when this method was carried out on the alcohol **40**, the yield of the alkene derivative **41** was low (24%), with substantial quantities of un-reacted alcohol (Scheme 2.4), even with the continued heating for 48 h at 70 ºC.



40 **41**

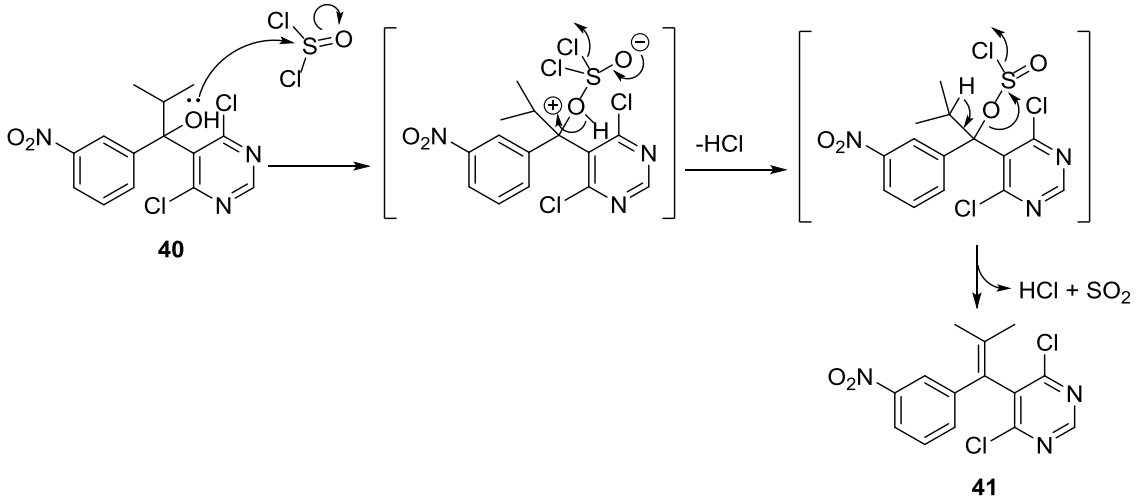
Scheme 2.4 Dehydration reaction of the alcohol **40** to the alkene **41**, using $SOCl_2$ and P_2O_5 in CH_2CH_2.

The reaction was then performed using neat pre-distilled $SOCl_2$, by heating the alcohol **40** with $SOCl_2$ at reflux under a N_2 atmosphere for 48 h (Scheme 2.5). After basic aqueous workup, the alkene **41** was isolated in 70% yield. Analysis of the ^1H NMR spectrum of the alkene **41** showed the disappearnce of two peaks at 3.38 ppm and 3.64-3.58 ppm from **40**, indicating the loss of the alcohol functionality. Analysis of the ^{13}C NMR spectra showed a shift of the propanol C1 of **40** (81.7 ppm) to 140.1 ppm for the propene C1 in **41**, indicating the formation of the double bond.



Scheme 2.5 Dehydration of the alcohol **40** using neat $SOCl_2$.

The proposed dehydration mechanism of the reaction begins with the nucleophilic attack of the alcohol to the sulfonylchloride. Elimination of HCl results in the chlorosulfite intermediate which eliminates to yield SO_2 and a second mole of HCl and the alkene **41** as shown in Scheme 2.6.

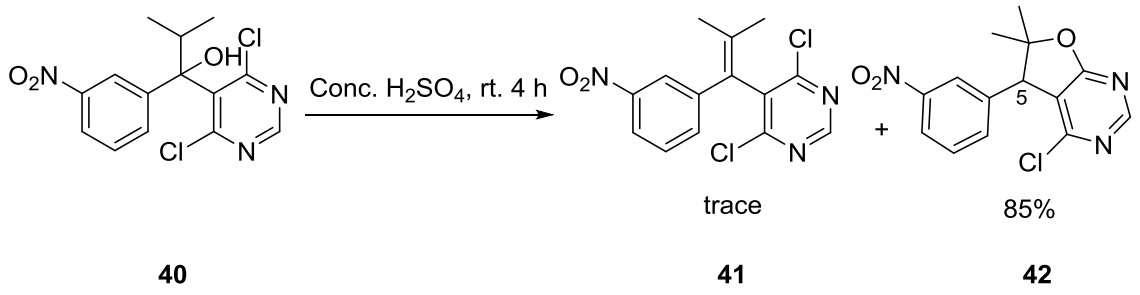


Scheme 2.6 The proposed dehydration mechanism of the alcohol **40** to form **41**.

The difference in the yield of the alkene **41** obtained in this work (70%) from the yield obtained by the previous other derivatives[156,157] is attributed to the difference in the electronic environment of the aromatic ring carrying the deactivating nitro group **40**. The electron withdrawing nitro group on the alcohol **40** may be responsible for slowing the rate of the reaction (Scheme 2.6), especially the last elimination step of the HCl and SO_2 in the proposed mechanism.

The yield of the alkene **41** using $SOCl_2$ dehydration mechanism was satisfactory, however, the reaction required 48 h. Therefore, an alternative dehydration

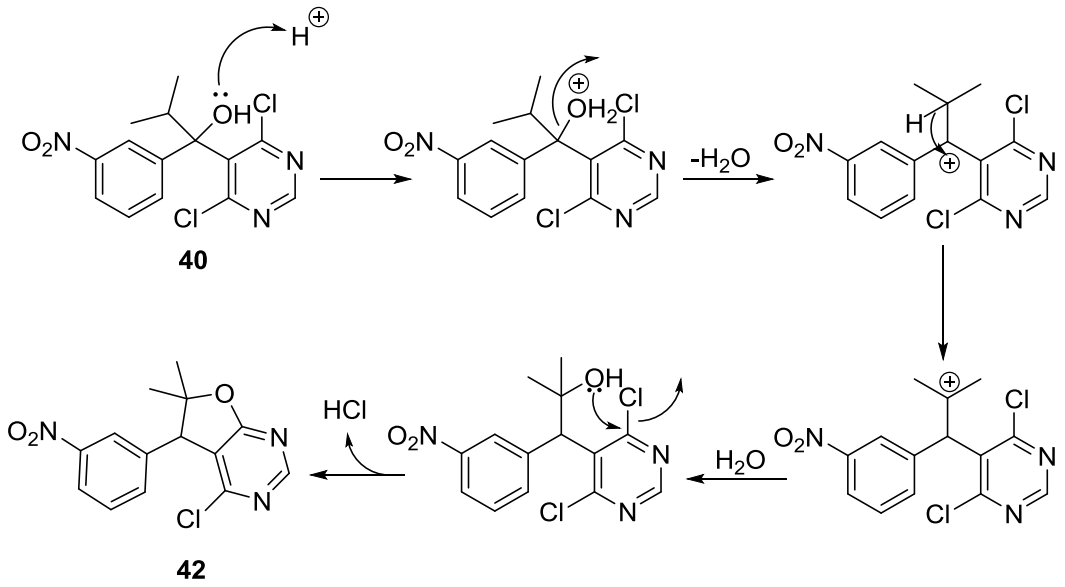
method was investigated. Conc. sulphuric acid is a well know dehydrating reagent for primary alcohols *via* an E2 mechanism, and secondary and tertiary alcohols *via* an E1 mechanism, to give the corresponding alkenes in both cases.[159] This sulphuric acid dehydration reaction was performed on the alcohol **40** by adding the later, portion-wise, to conc. H_2SO_4 and allowing the mixture to stir at room temperature (Scheme 2.7). The reaction was allowed to stir until all the alcohol **40** was consumed. After workup, the alkene **41** spot was found to be present in trace (not changing) compared to the major product, which was found to be 4-chloro-6,6-dimethyl-5-(3-nitrophenyl)-5,6-dihydrofuro[2,3-*d*]pyrimidine **42** derivative (Scheme 2.7). Analysis of the ^{1}H NMR spectrum showed a singlet peak resonating at 4.55 ppm integrating for one proton, which was assigned to H5 of the furo-pyrimidine skeleton. Analysis of the HRMS spectrum showed a peak at 306.0648 *m/z* which was assigned to the molecular formula $C_{14}H_{13}ClN_3O_3$ (MH+), indicating that the alcohol oxygen atom was still present in the structure



Scheme 2.7 Dehydration reaction of the alcohol **40** using H_2SO_4, producing the furo-pyrimidine **42**.

The proposed mechanism for the formation of the furo[2,3-*d*]pyrimidine **42** was attributed to the rearrangement of the carbocation produced after the E1 elimination step (Scheme 2.8), where the positive charge is transferred to carbon 2 of the propyl chain (through a hydride shift). The planarity of the benzylic carbocation with the benzene and the pyrimidine ring might be responsible for such a rearrangement. This later carbocation is further attacked by water giving another tertiary alcohol on C2 of the propyl chain. S_N2 attack of the OH group to Cl gave the cyclised product **42**. This reaction (Scheme 2.7) is considered as a new convenient two step synthesis of the

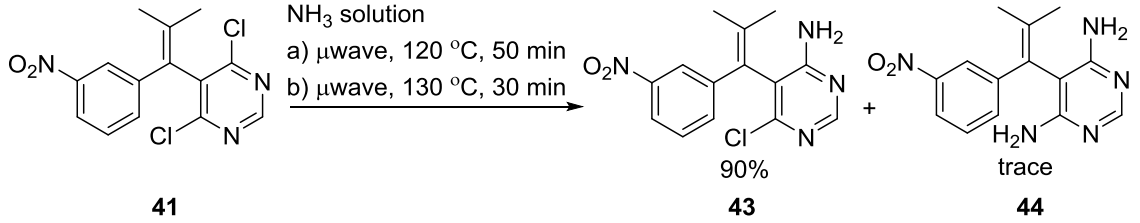
furo[2,3-*d*]pyrimidine heterocycle, which is rarely reported in the literature. Existing methods for the synthesis of this scaffold require more harsh synthetic[160,161] conditions.



Scheme 2.8 The proposed mechanism of the formation of furo[2,3-*d*]pyrimidine derivative **42**.

2.2.4. *Mono-amination of the alkene 41*

The nucleophilic substitution of ammonia onto the pyrimidine ring of the alkene **41** has been previously optimized on similar scaffolds using a microwave assisted reaction.[157] Therefore, the alkene **41** was suspended in ammonia solution (25%), and the mixture heated and irradiated in a microwave reactor at two stages; the first stage involved heating the mixture at 120 ºC for 50 min, followed by a slight increase in the temperature to 130 ºC, for a 30 min (Scheme 2.9). These conditions maximized the production of the mono-amino substitution product **43** and minimized the production of the di-amino product **44**. Analysis of the ^{1}H NMR spectrum of **43** showed a broad peak at 5.27 ppm, which was assigned to the NH_2 group. Analysis of the ESI-MS spectrum showed a peak at 307 *m/z* for the molecular ion (M+, ^{37}Cl), indicating the replacement of one Cl atom in the alkene **41** by an amino group.



Scheme 2.9 Selective amination of the dichloro-pyrimidine **41**.

The reaction proceeded *via* a S_NAr mechanism where the lone pair of electrons of ammonia attacked C4 of the pyrimidine ring followed by rearomatisation of the ring displacing the chlorine atom resulting in the formation of **43**.

2.2.5. *Final amination of 43 for the final AAPM derivative*

The final amination of the amino alkenes had been a challenge in the synthesis of this series, in this work as well as the previous work.[157] Only the lead compound, the chloro-AAPM **27**, and the unsubstituted AAPM **34** were synthesized previously in our group, with multiple attempts to generate more derivatives.[156,157]

The previous synthesis of the AAPM **27** utilised the reaction between the racemic, highly viscous, propylamine side-chain **33** in a solvent free reaction with the aminochloropyrimidine **32** to give the desired AAPM **27** (Figure 2.3) in 40% yield.[155,156]

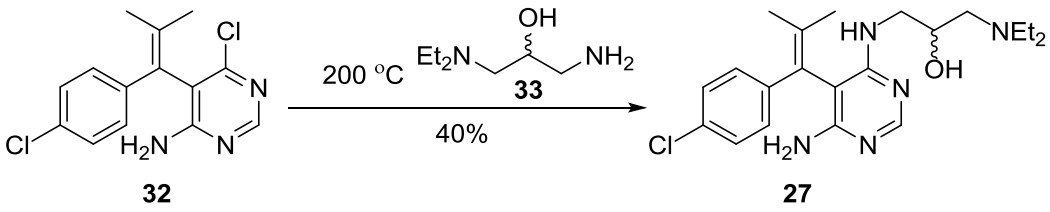


Figure 2.3 The synthesis of the AAPM **27** from the aminochloropyrimidine **32**.

When this synthetic strategy was used previously in our group for the synthesis of the unsubstituted AAPM **34**, the reaction failed with no evidence for the formation of the AAPM **34**, even after conventional heating of the reactants at 200 ºC for 5 h (Figure 2.5).[157] Instead, further investigation of the reaction mixture indicated the

polymerization of the poly aminic side chain **33**.[157] However, access to the AAPM **34** was achieved through an optimization process previously investigated in our group,[157] using a microwave irradiation method (Figure 2.4), with a poor yield (21%).[157]

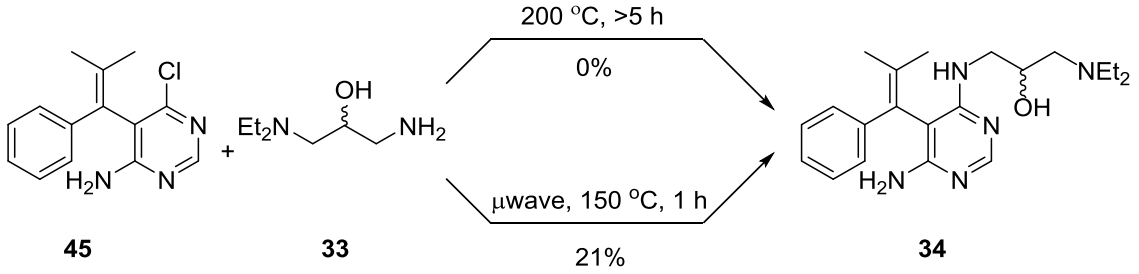
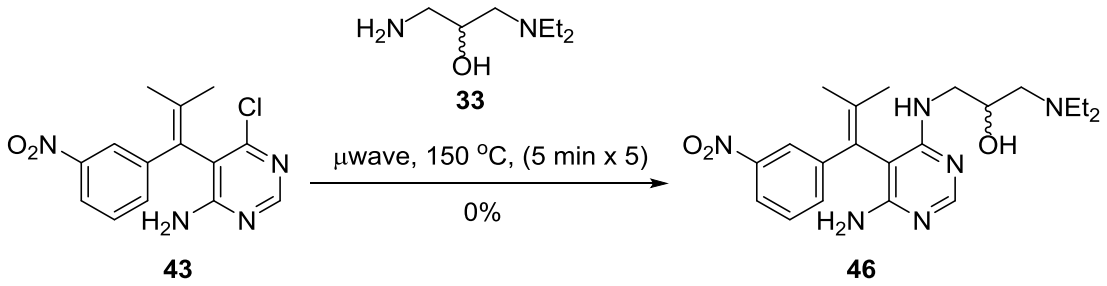


Figure 2.4 Reactions of the aminopyrimidine **45** with the amine **33** for the synthesis of the AAPM **34**.

Therefore, the aminochloropyrimidine **43** was heated with the amine **33** under microwave irradiation at 150 ºC for 5 min (Scheme 2.10). The reaction cycle was repeated 5 times until the complete disappearance of the starting aminochloropyrimidine **43** as indicated by the TLC analysis (90% ethyl acetate in petroleum spirit). Further analysis of the reaction revealed the presence of polymeric baseline material and the absence of material resembling the AAPM **46**. This indicated the disintegration of the heterocyclic structure. Although harsh conditions are required to overcome the steric barrier of the reaction, prolonged periods of heating result in the destruction of the material. Therefore, lower temperature were investigated, Table 2.1 summarises the series of microwave reactions attempted using the aminochloropyrimidine **43** and the amine **33**.



Scheme 2.10 Attempted synthesis of the AAPM **46** using the microwave irradiation.

Table 2.1 Attempted microwave reactions of **43** with **33**.

Entry	Microwave temp. (ºC)	Reaction time (min)	Yield (%) of **46**
1	120	50	0
2	130	30	0
3	140	40	0

For all entries (Table 2.1), TLC analysis was performed to gauge the outcome of each reaction. For entry 1, the staring material **43** was the major spot, with the additional appearance of the polymeric baseline material. Mass spectrometric analysis of the crude reaction mixture did not show any peaks that could be assigned to the formation of the AAPM **46**. Increasing the temperature to 130 ºC (entry 2) and 140 ºC (entry 3) did not produce any differences from the results from entry 1, with the exception of increased decomposition.

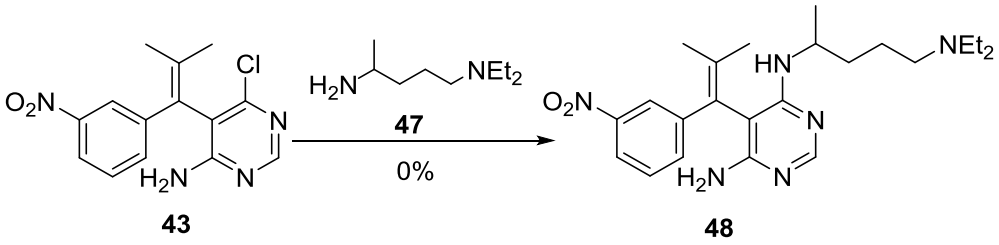
The next attempt to synthesise **46** used solvent in order to solubilise the highly viscous amino side chain **33**, and for the ease of recovery of the aminochloropyrimidine **43**. In the previous optimized process of the synthesis of AAPM **34**, dichlorobenzene (DCB) was used, however, it was demonstrated not to improve the reaction, nor the recovery of the un-reacted starting material.[157] Therefore, the reaction was attempted here using water as a solvent, because of the amino side chain **33** solubility in water reduced the high viscosity of the amine **33** allowing a greater chance of reaction with the aminochloropyrimidine **43**. Water was also selected as it had been successfully used for many microwave assisted reactions.[162] Therefore, the aminochloropyrimidine **43** was heated with amino side chain **33** in a microwave tube at 100 ºC for 5 min. TLC analysis (90% ethyl acetate in petroleum spirit) indicated the presence of the starting materials with polymeric baseline material. The temperature was increased to 150 ºC and the reaction was held at this temperature for 10 min. TLC analysis after the 10 min at 150 ºC indicated that the concentration of the starting material was decreasing and the concentration of the baseline material increased. Furthermore, low resolution mass spectrometric analysis did not show any evidence indicating the formation of the desired AAPM **46**. The aminochloropyrimidine **43** was recovered in 60%.

The difficulties in the synthesis of the AAPM derivatives along with the previous attempts[157] were attributed to the polymerization of the amino side chain **33** at

the reaction temperatures. The electronic effect of the substituents on the phenyl ring (Cl in AAPM **27**, H in AAPM **34** and NO_2 in the unformed AAPM **46**) is also believed to play a role in the amination reaction, where the Cl atom might activate the amination and the NO_2 group may deactivate the amination reaction. No investigations were performed at this point to confirm this hypothesis either in the previous work or in this work.*

2.3. Modification in the amine side chain of the AAPM derivatives

An alternative amine side chain, N^1,N^1-diethylpentane-1,4-diamine **47**† was selected as a replacement for the polymeric side chain **33**, as it is identical to the amino side chain of the antimalarial chloroquine **8**. Furthermore, the amine **47** is a liquid with low viscosity compared to the amine **33**. Therefore, it was thought that it would react more easily with the aminochloropyrimidine **43**. The later was reacted with amine **47** in different conditions (Scheme 2.11) using either conventional heating or microwave irradiation (Table 2.2).



Scheme 2.11 Attempted reaction of the aminochloropyrimidine **43** with the amine **47**.

* Personal communication with Andrew Stevens (ref 157), who synthesized the AAPM **34**, and attempted in other AAPM derivatives.

† Was purchased from Aldrich and further distilled under N_2 at 70 ºC to provide a colourless liquid.

Table 2.2 Attempted reaction conditions for the synthesis of the target compound **48**.

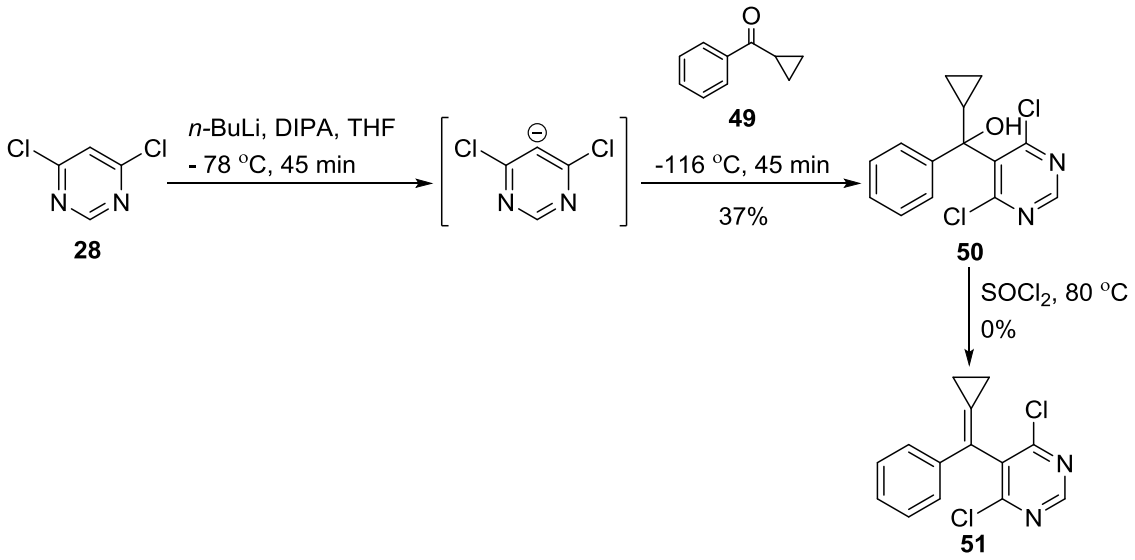
Entry	Heating method	Reaction temp. (ºC)	Reaction time	Yield (%) of **48**
1	Microwave	150	3 min	0
2	Microwave	150	10 min	0
3	Microwave	170	5 min	Trace
4	Microwave	200	3 min	Trace
5	Microwave	200	12 min	0
6	Conventional	150	12 h	0
7	Conventional	200	48 h	0

From Table 2.2, the first attempt of the reaction (Scheme 2.11) used microwave irradiation (entry 1); this method resulted in no change as indicated by TLC and mass spectrometric analyses. Increasing the reaction time to 10 min (entry 2) did not produce any change in the reaction. The temperature was then increased to 170 ºC (entry 3) and the reaction was held for 5 min. The mass spectrum of the reaction mixture showed a weak peak at 427 *m/z*, indicating a trace of the product **48**. Increasing the temperature to 200 ºC for 3 min (entry 4) to push the reaction to completion did not improve the production of **48** but still the weak 427 *m/z* peak in the mass spectrum could be observed. When the hold time was increased to 12 min (entry 5), the weak 427 *m/z* peak of the product **48** could not be seen indicating the decomposition or polymerization of the product. In entries 1-4, 70-73% of the starting material was recovered indicating that little reaction had taken place, while in entry 5, no starting material was recovered from the reaction mixture. The reaction (entry 4) was then repeated on fresh starting materials, the content was then subjected to preparative TLC (PTLC) using 10% water in acetonitrile, the baseline band was isolated. Analysis of the mass spectrum of this band revealed the presence of the **48** peak at 427 *m/z* and the amine peak **47** at 158 *m/z*. However, analysis of the ^{1}H NMR spectrum of this fraction showed mainly complex polymeric protons in the aliphatic region (0.6-2.5 ppm) and the compound **48** was not separated.

When the reactants were heated using the conventional heating method (entries 6, 7), no evidence for the formation of **48** was detected by either mass spectrometric or NMR analysis, and the starting aminochloropyrimidine **43** was completely decomposed.

2.4. Decreasing the steric hindrance of the dimethyl groups

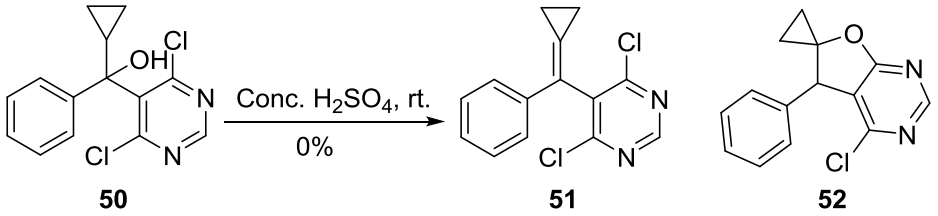
The poor reactivity of the aminochloropyrimidine **43** with either the amine **33** (Scheme 2.10) or the amine **47** (Scheme 2.11) argued for the difficulties in the reaction being due to the steric clashes with the dimethyl group. Therefore, the dimethyl groups in **43** were replaced with a cyclopropyl group, in an attempt to give the aminochloropyrimidine **51**. The cyclopropyl group was selected in order to tether the dimethyl groups together and minimize the steric hindrance. The attempted access to **51** was performed in a similar synthetic pathway (Scheme 2.12) as that used for **41**, using cyclopropylphenyl ketone **49**.



Scheme 2.12 Attempted synthesis of the alkene **51** through the alcohol **50**.

The alcohol **50** was isolated in 37% yield compared to 50% of the alcohol **40**. Analysis of the ^{1}H NMR spectrum of **50** showed a broad peak at 2.83 ppm, assigned to the OH substituent. The ESI-MS spectrum showed a peak at 295.1 *m/z* assigned to the (M+, ^{35}Cl) of the alcohol **50**. However, the alcohol **50** was found to be unstable, with decomposition occurring immediately after column chromatography (80% ethyl acetate in petroleum spirit), as indicated by TLC analysis. These unstable by-products could not

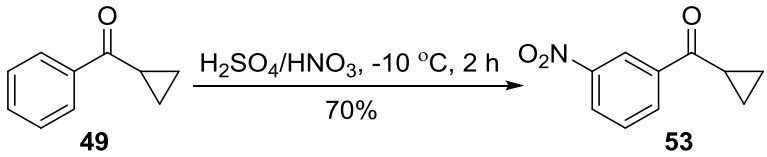
be separated and identified. Therefore, the dehydration reaction was carried out immediately after the isolation of **50** from the column. Stirring the alcohol **51** with $SOCl_2$ at room temperature, or at 80 ºC, caused only charring of the alcohol **51**. The dehydration reaction using conc. H_2SO_4, that was carried out on alcohol **40** (Scheme 2.7), was also attempted on alcohol **50** to investigate the possibility of forming the alkene **51**, or the furopyrimidine **52** product (Scheme 2.13). However, charring was also observed after 5 min of stirring, with no formation of **51** or **52**.



50 **51** **52**

Scheme 2.13 Attempted conc. H_2SO_4 reaction on the alcohol **50**.

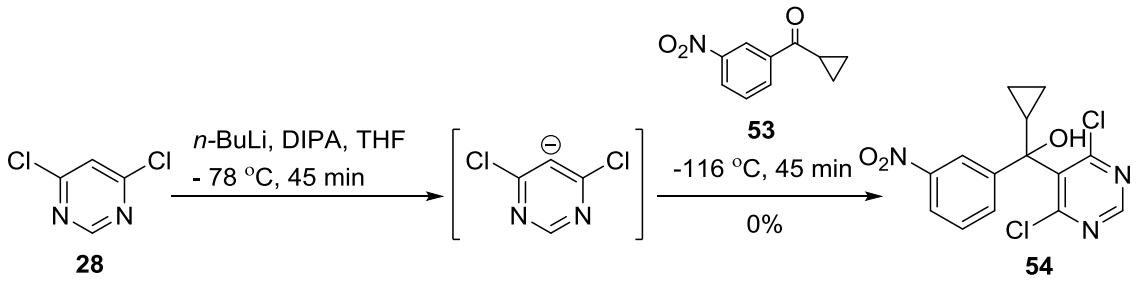
Compared to alcohol **40**, alcohol **50** does not contain a nitro group in the phenyl ring. The next investigation was to check whether the presence of the nitro group on the phenyl ring of **50** would stabilize the alcohol or not. Therefore, the cyclopropylphenyl ketone **49** was first nitrated using conc. H_2SO_4/HNO_3 at -10 ºC to afford the cyclopropyl(3-nitrophenyl)methanone **53** in 70% yield (Scheme 2.14).



49 **53**

Scheme 2.14 Nitration of cyclopropylphenyl ketone **49**.

The phenone **53** was then subjected to the LDA addition reaction with the 4,6-dichloropyrimidine **28** (Scheme 2.15). Despite subjecting the substrates to identical reaction conditions previously used for the synthesis of the alcohols **40** and **50**, the alcohol **54** was not formed and the starting phenone **53** was recovered (50%). The percentage consumed from the phenone **53** indicated that the reaction may have occurred but the product was unstable, and immediately decomposed.



Scheme 2.14 Attempted LDA addition reaction of the phenone **53** with the 4,6-dichloropyrimidine **28**.

These investigations into the synthesis the AAPM **46**, compound **48** and the alcohol **54**, as well as the instability of the alcohol **50**, argued for the importance of decreasing the steric hindrance around the Cl atom to be substituted with the amine, and also argued for the need of further simplification of the AAPM structure.

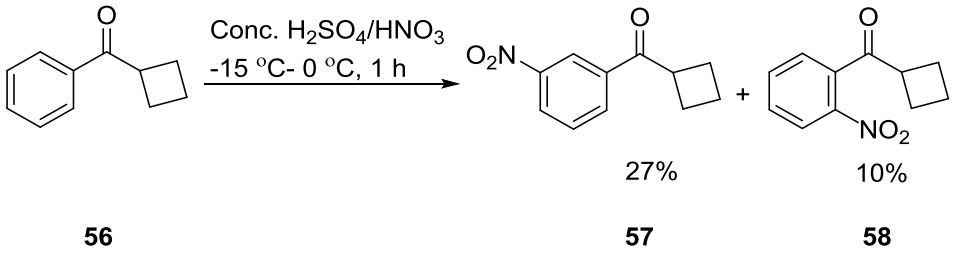
2.5. Replacing the 4,6-dichloropyrimidine with 4-chloropyridine

One strategy to simplify the core structure of the AAPM precursors involved replacing the bulky 4,6-dichloropyrimidine **28** (Scheme 2.2) with the simpler 4-chloropyridine.HCl **55**. Lithiation of the *meta* position of 4-chloropyridine **55**, with subsequent addition of aldehydes has been previously reported,[163] where the electrophilic addition was performed at -78 ºC without the need to cool the reaction to -116 ºC.[163] This change of the addition temperature would be advantageous, compared to the example of 4,6-dichloropyrimidine **28**, in avoiding the difficulty of maintaining the reaction temperature at -116 ºC. During this modification, the nitro group was kept on the phenyl rings of the reacting phenones, ketones **38**, **53** and the nitrocyclobutylphenyl ketone **57** were investigated in this modification step.

2.5.1. *Synthesis of the nitrocyclobutylphenyl ketone 57*

Ketone **57** was accessed through the nitration of the precursor phenone **56** using a nitrating mixture of conc. H_2SO_4/HNO_3 at -15 ºC (Scheme 2.15). After completion of the reaction and the aqueous workup, the oil left was subjected to column chromatography using 30% ether in petroleum spirit. The desired *meta* isomer

cyclobutyl(3-nitrophenyl)methanone **57** was isolated in 27% yield, and the *ortho* isomer cyclobutyl(2-nitrophenyl)methanone **58** in was isolated 10% yield.



Scheme 2.15 Nitration of the cyclobutylphenyl ketone **56**.

The *ortho* nitration of ketones carrying bulky groups has been previously observed as in case of adamantyl bearing aromatic ketones.[164] Therefore, in the presence of the bulky cyclobutyl ring, the formation of the *ortho* derivative **58** was also possible, and could be explained by formation of the nitronium or proton complex with the carbonyl of cyclobutylphenyl ketone **56** and establishment of pre-equilibrium included both this form and its mutual competition. The increase of the nitronium-carbonyl complex concentration leads to an increase of the amount of *ortho* nitro regioisomer **58** (Figure 2.5) by the nucleophilic attack of the benzene ring from the *ortho* position. Optimization of this reaction (Scheme 2.15) could be investigated by changing the reaction conditions, however, this was out of the scope in this work.

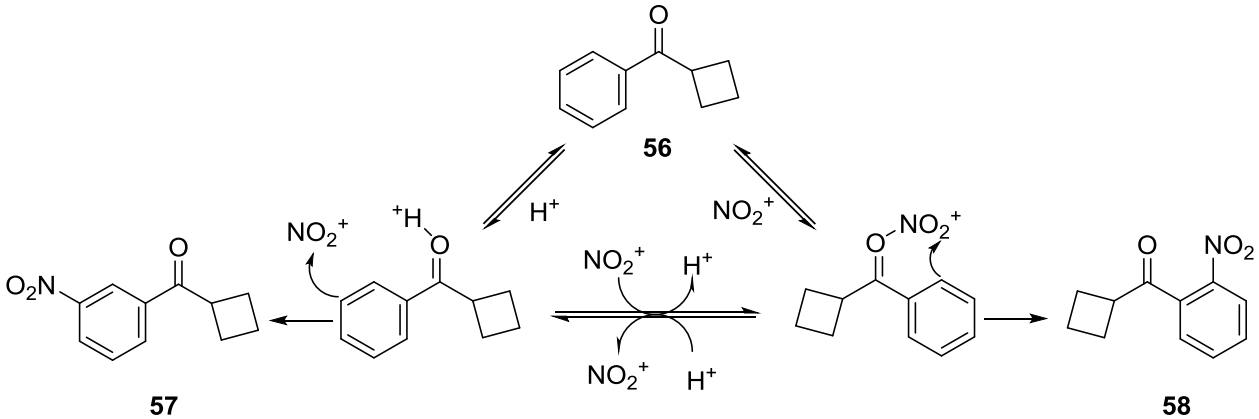
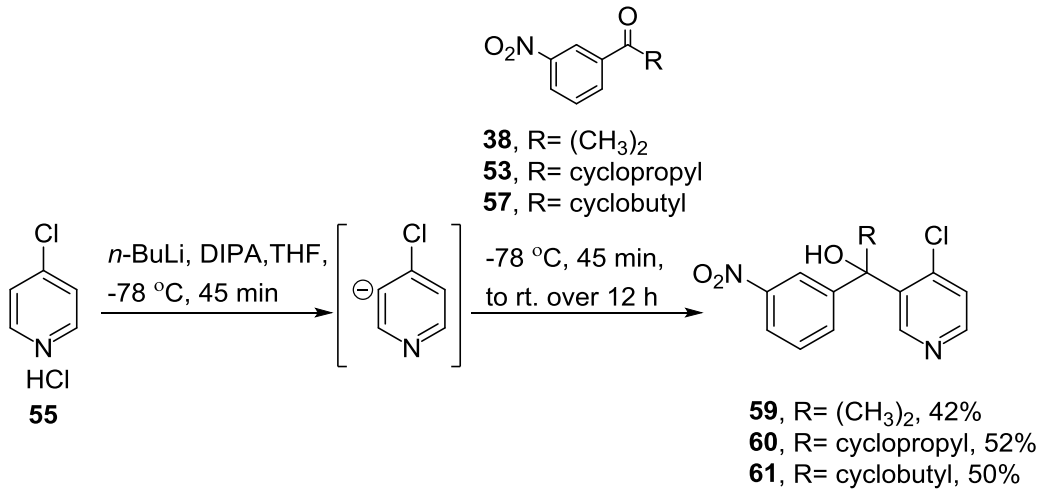


Figure 2.5 Proposed mechanism for the nitration of **56** and the formation of **58**.[164]

2.5.2. *Addition reactions of ketones to the 4-chloropyridine.HCl 55*

The first step in the addition reaction was the lithiation of the 4-chloropyridine **55** using a freshly prepared LDA solution at -78 ºC as previously discussed (Scheme 2.2). The prepared LDA solution was then transferred to the solid 4-chloropyridine HCl **55**, and the solution was left to react at -78 ºC and to form the lithiated pyridinyl anion (Scheme 2.16). The completion of the reaction was indicated by the disappearance of the solid 4-chloropyridine HCl **55**, which took almost 45 min. To this anion solution, was added dropwise a solution of the appropriate ketone (**38**, **53** or **57**) in dry THF at -78 ºC. The reaction temperature was kept at -78 ºC for 45 min, after which the reaction mixture was allowed to warm up to room temperature over 12 h. After quenching the reaction and an acidic work up, the alcohols **59-61** were obtained after column chromatography in 48-52% yield. The same dry reaction conditions discussed previously (Scheme 2.2) were also utilized. Addition of the ketones at -78 ºC (Scheme 2.16) rather than at -116 ºC (Scheme 2.2), was advantageous and could be more easily controlled. Moreover, the acidic workup conditions in case of the pyridine based alcohols **59-61** facilitated the isolation of relatively pure products. This was indicated by the improved yield and less side products observed by TLC analysis, compared to the results obtained before when using pyrimidine (Scheme 2.2). Analysis of the ^{1}H NMR spectrum of alcohol **59** showed a broad peak at 3.25 ppm assigned to the OH substituent, whereas the EI-MS spectrum showed a peak at 306 *m/z* which was assigned to the (M+, ^{35}Cl).

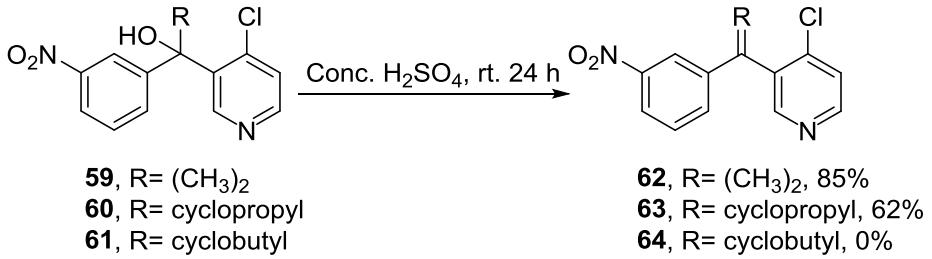


Scheme 2.16 Addition of the ketones (**38**, **53**, **57**) to the 4-chloropyridine HCl **55**.

The reaction was attempted using 2 molar equivalents of the prepared LDA solution, however, this was insufficient to react with the 1 molar equivalent of 4-chloropyridine HCl **55** and some of the solid remained undissolved. On addition of the electrophile (ketones **38**, **53** and **57**), the reaction proceeded, however, the isolated yields were lower than those obtained in Scheme 2.16, due to the presence of a lower concentration of the lithiated pyridinyl anion (Scheme 2.16). Using 2.5 molar equivellents of LDA was found to completely react with 4-chloropyridine HCl **55**.

2.5.3. *Dehydration of the alcohols 59-61*

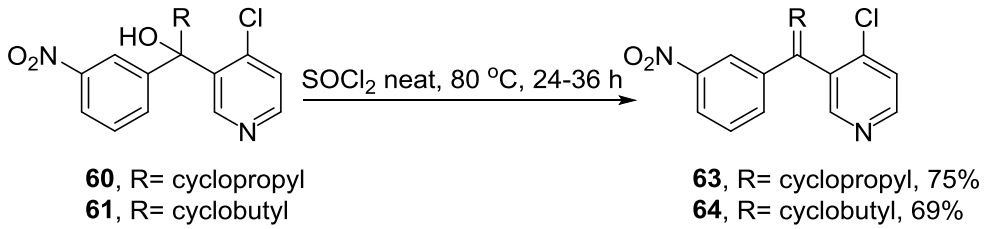
The first attempted dehydration reaction used the conc. H_2SO_4 method. Therefore, a mixture of the appropriate alcohol (**59-61**) and conc. H_2SO_4 was stirred at room temperature for 24 h and the reaction was monitored by TLC analysis at 1 h time intervals (Scheme 2.17). Although the reactions were clean and complete by TLC analysis, in the case of the alcohol **59** and the cyclopropyl alcohol **60**, the cyclobutyl alcohol **61** underwent charring after 30 min of stirring with conc. H_2SO_4. The dimethyl alkene analogue **62** and the cyclopropyl alkene **63** were isolated after aqueous basic workup conditions as yellow oils in 85% and 62% yield, respectively. Analysis of the ^{1}H NMR spectrum of the alkene **62** showed the disappearance of the OH substituent of the alcohol that was resonating at 3.25 ppm, and the adjacent proton that was resonating as a multiplet peak at 3.12-3.09 ppm in **59**, indicating the formation of the double bond. Analysis of the ^{13}C NMR spectrum showed a shift of the propanol C1 in **59** (79.7 ppm) to 138.5 ppm for the propene C1 of **62** .Analysis of the EI-MS spectrum of **62** showed a peak at 288 *m/z*, which was assigned to M+.



Scheme 2.17 Dehydration of the alcohols (**59-61**) using conc. H_2SO_4.

Upon the workup of the charred solution in case the alcohol **61**, neither the alcohol nor the alkene **64** were detected, instead, a polymeric dark line was present by

TLC analysis. Therefore, the $SOCl_2$ dehydration method was attempted, where the alcohols **60** or **61** were added to neat $SOCl_2$, and the solutions were heated at 80 ºC under N_2 atmosphere for 24 h and 36 h, respectively (Scheme 2.18). Upon cooling the reactions, ice was added, and the solutions were neutralized with 2 M KOH. After CH_2CH_2 extraction and column chromatography, the oily alkenes **63** and **64** were isolated in 75% and 69% yield, respectively. The yield of the alkene **63** was slightly improved over using the conc. H_2SO_4 method (Scheme 2.17).



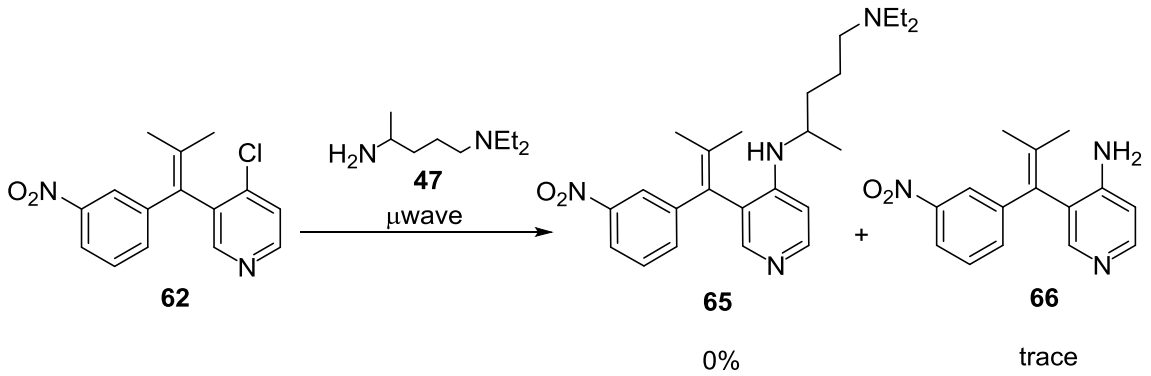
60, R= cyclopropyl
61, R= cyclobutyl

63, R= cyclopropyl, 75%
64, R= cyclobutyl, 69%

Scheme 2.18 Dehydration of the alcohols **60** and **61** using neat $SOCl_2$.

2.5.4. *Final amination of the alkenes 62-64*

The amine **33** was avoided in this trial based on the results previously discussed (Scheme 2.10). The amine **47** was selected for the amination reaction on the alkenes **62-64** and a series of reactions, using different conditions, were carried out on the alkenes **62-64** depending on the availability of each. The alkene **62** was mixed neat with the amine **47** and the mixture was placed in a microwave tube, followed by irradiation under different conditions (Scheme 2.19, Table 2.3).



62

65

0%

66

trace

Scheme 2.19 Attempted microwave reaction of the alkene **62** with the amine **47**.

Table 2.3 Microwave reaction conditions of the alkene **62** with the amine **47**.

Entry	Microwave temp. (ºC)	Reaction time (min)	Yield (%) of **65**	Yield (%) of **66**
1	135	10	0	0
2	150	10	0	0
3	170	20	0	0
4	200	5	0	0
5	250	20	0	0
6	280	20	0	trace
7	280	60	0	trace

After each entry (1-7, Table 2.3), a sample of the reaction mixture was taken for TLC and low resolution mass analysis. In entries 1-5, TLC analysis indicated no change in the reactants. Furthermore, the analysis of EI-MS spectra of the samples showed peaks at 288 *m/z* (M+ of **62**) and the amine **47** peak at 158 *m/z*. The stability of the alkene **62** at 250 ºC was encouraging to increase the temperature to 280 ºC. When the alkene **62** was heated with the amine **47** at 280 ºC for 20 min (entry 6), TLC analysis indicated no change, however, EI-MS spectrum showed a new peak at 269 *m/z*, which corresponded to the (M+) of the amine **66**. When the reaction was heated further at 280 ºC for 60 min (entry 7), the alkene **62** disappeared and the the peak at 269 *m/z* (amine **66**) was still present in the EI-MS spectrum. TLC analysis of the reaction mixture of entry 7 presented a polymeric baseline long spot and the amine **66** was not isolated. Formation of the amine **66** at high temperature not only indicated the decomposition of the amine **47** releasing the NH_2 nucleophile, but also indicated the weak reactivity (electrophilicity) of the alkene **62** towards the amine **47**. This lack of reactivity towards **47** and the reactivity with the released nucleophile NH_2, can be explained based on steric hindrance with the amine **47** being hindered from replacing the 4-chloro on the pyridine ring of **62**.

The instability of the amine **47** and the lack of reactivity of the alkene **62** under the normal S_NAr conditions required a different strategy. Therefore, the use of a catalyst such as palladium [Pd(0)], in mild conditions was attempted to facilitate the reaction and to avoid the degradation of the amine **47**. The amine **47** was reported to react with

4-chloroquinoline **67** in the presence of palladium acetate (4 mol%), DPEphos* (8 mol%) and potassium phosphate (K_3PO_4, 2.5 equivalent) and the mixture was heated in 1,4-dioxane for 18 h at 85 ºC, where the final product **68** was isolated in 93% yield (Figure 2.6).[165]

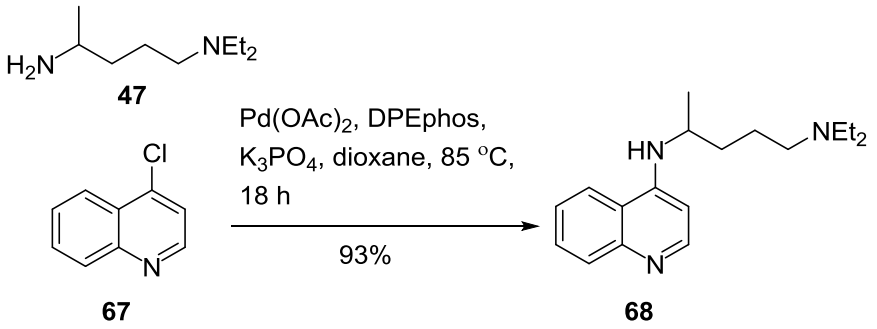
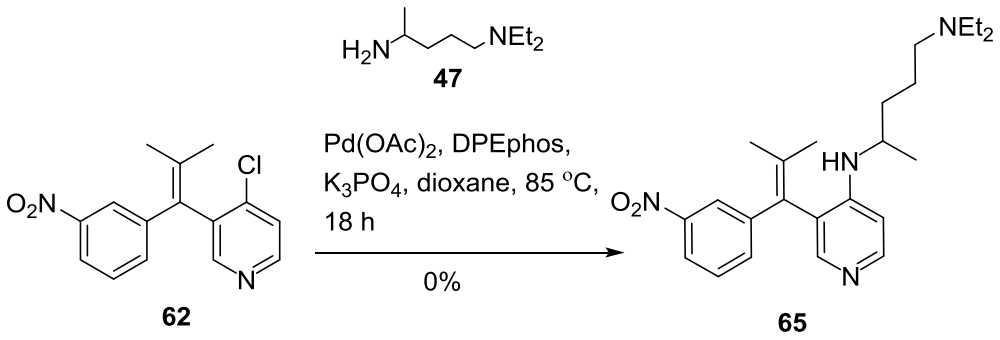


Figure 2.6 Reported amination[165] reaction using the amine **47** with 4-chloroquinoline **67** using a Pd catalyst.

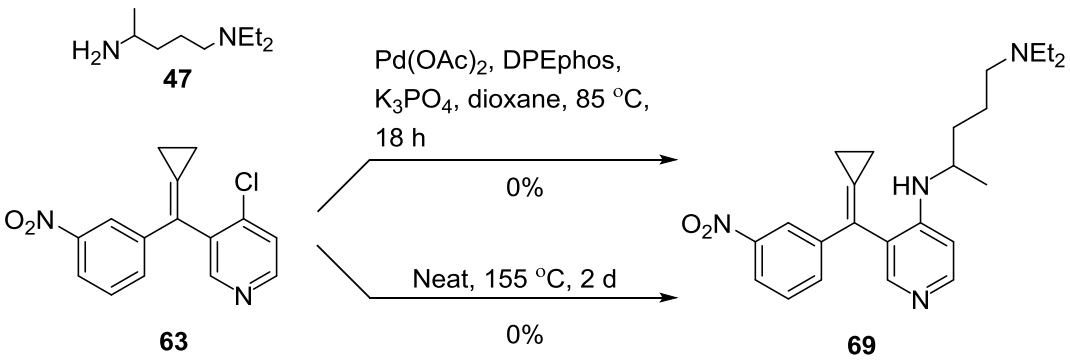
The presence of the chlorine atom in position 4 of the quinoline **67**, makes it analogous to the case in the alkene **62**, however, there is no steric hindrance around the chlorine atom in **67**. Therefore, the alkene **62** was placed with the amine **47** in a sealed tube, to which was added palladium acetate (4 mol%), DPEphos (8 mol%) and K_3PO_4 (2.5 equivalent), and the tube was degassed for 10 min using N_2, the tube was sealed and then heated at 18 h at 85 ºC (Scheme 2.20). Upon TLC analysis of the mixture, no change was observed. Therefore, the content of the tube was transferred to a microwave tube and the reaction was irradiated with microwave at 100-160 ºC for 10 min to 1 h. No change in the reaction was detected until reaching 160 ºC for 1 h, where the reactants decomposed as indicated by the TLC analysis where several minor spots (>7) appeared that were difficult to separate.

* (oxydi-2,1-phenylene)bis(diphenylphosphine).



Scheme 2.20 Attempted Pd catalysed amination of the alkene **62** using the amine **47**.

The failure of the above reactions (Schemes 2.19, 2.20) argued for the difficulty of the reactions in the presence of the steric hindrance induced mainly by the methyl groups in both **62** and **47**. Therefore, the less hindered, alkene **63**, with the cyclopropyl chain instead of the two methyl groups in **62**, was reacted with the amine **47** using the same Pd catalysed conditions (Scheme 2.21). At the same time, the alkene **63** was reacted with the amine **47** using a conventional heating method in a sealed tube, to test the reactivity without the Pd catalyst (Scheme 2.21).



Scheme 2.21 Attempted reactions of the cyclopropyl alkene **63** with the amine **47**, using the Pd catalysed amination and the free S_NAr reaction.

Using the Pd catalysed conditions, the product **69** was not detected and only the alkene **63** was recovered. Using the thermal heating method, the reaction was heated until the starting material disappeared. No characteristic peaks were identified in the EI-MS/ESI-MS spectra, therefore, the excess amine **47** was then removed by distillation

under vacuum at 120 ºC. Further analysis of the EI-MS and ^{1}H NMR spectra of the oily residue did not provide any evidence indicating the formation of **69**, instead, a polymeric material was detected that had complex multiplicity in the aliphatic regions (1-3 ppm) with no aromatic protons.

Another S_NAr reaction of the amine **47** with 9-chloroacridine **70** was reported in 2009, utilizing phenol as a solvent and heating the reaction at 100 ºC for 15 min, with the product **71** isolated in 65% yield[166] (Figure 2.7).

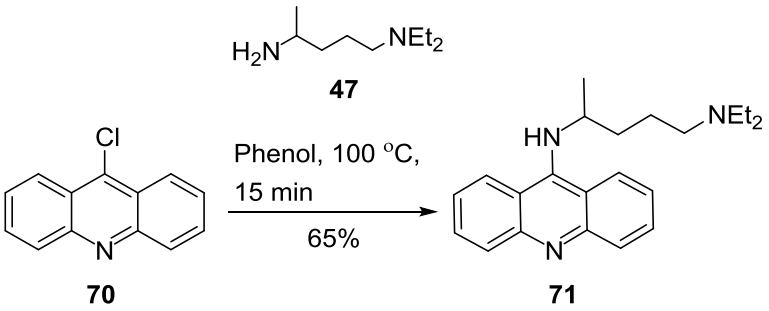
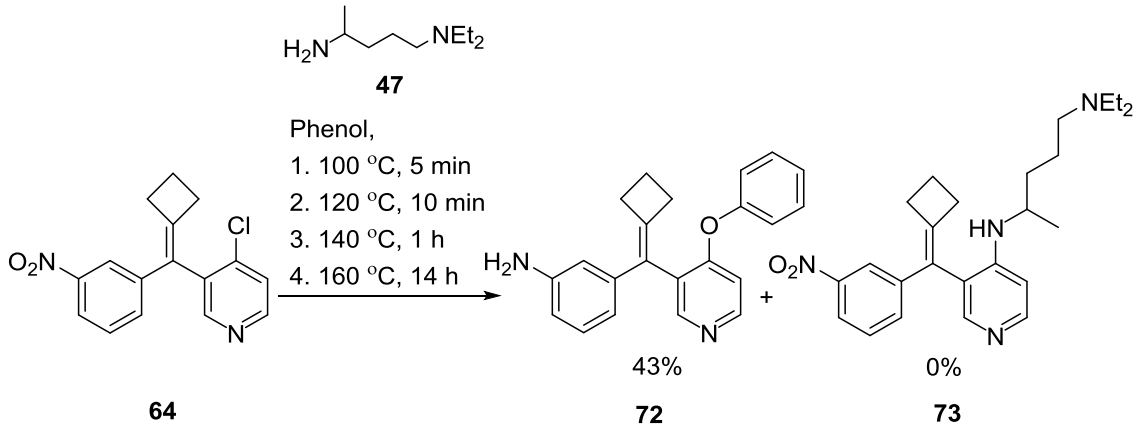


Figure 2.7 Reported S_NAr reaction of the 9-chloroacridine **70** with the amine **47** using phenol as a solvent.[166]

Therefore, the alkene **64** was placed in a sealed tube, phenol was added and then, the amine **47** was added. The tube was sealed and the reaction was heated at 100 ºC for 5 min (Scheme 2.22). A sample of the reaction was taken for TLC analysis, which indicated no change. The tube was then heated at 120 ºC for 10 min and 140 ºC for 1 h, after which, no change was observed in the reaction. Finally, the tube was heated at 160 ºC for 14 h. The tube was then cooled and the content was adsorbed onto silica and subjected to a flash column chromatography and eluted with 10% methanol in CH_2Cl_2 to give 3-(cyclobutylidene(4-phenoxypyridin-3-yl)methyl)aniline **72** as a colourless oil in 43% yield. Analysis of the EI-MS spectrum showed a peak at 328 *m/z* that was assigned to the M+. Analysis of the ^{1}H NMR spectrum showed a broad peak at 3.55 ppm which was assigned to the NH_2 group. The IR spectrum of **72** showed the disappearance of the NO_2 group absorption bands at 1527 and 1340 cm^{-1} from **64**, and showed an absorption band at 3256 cm^{-1} indicating the presence of an amino group, confirming the reduction of the nitro group on the phenyl ring of **64**.



Scheme 2.22 Reaction of the alkene **64** with the amine **47** using phenol as a solvent.

The nucleophilic substitution reaction of the chlorine atom by the phenol is a known reaction for the preparation of ethers (Ullmann ether synthesis), where the nucleophile (phenol) is heated with a chlorated substrate at elevated temperatures, in the presence of a base and copper metal, as in case of 4-phenoxypyridine **74**, which was prepared *via* reacting 4-chloropyridine HCl **55** with phenol in the presence of cesium carbonate (Cs_2CO_3), and Cu-source (Figure 2.8).[167] The same reaction can also occur in the absence of the copper metal, where the 4-chloropyridine HCl **55** could be heated with phenol at 150 ºC for 15 h (Figure 2.8).[168]

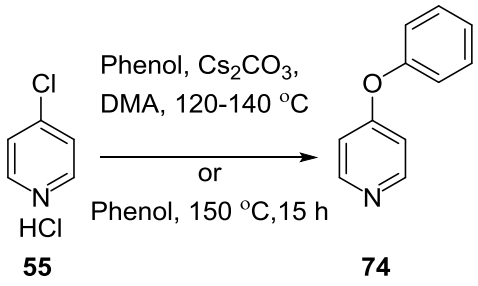


Figure 2.8 Reported reaction of 4-chloropyridine HCl **55** with phenol, forming 4-phenoxypyridine **74**.[167,168]

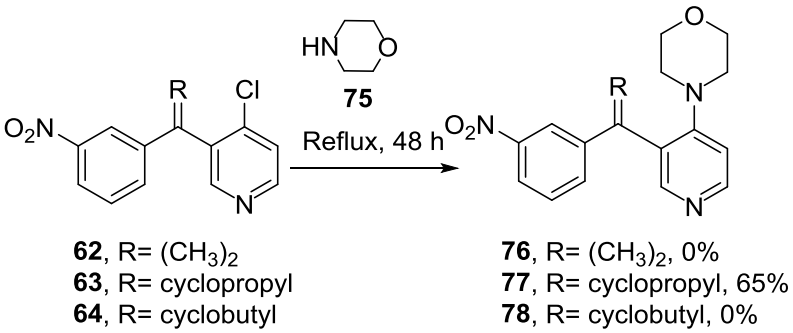
In our reaction (Scheme 2.22), two transformations are taking place, the solvent phenol first serves as a nucleophile and undergoes an S_N2 reaction, replacing the Cl

leaving group, producing the ether **72**. This was surprising because the phenol is a weaker nucleophile than the NH_2 group of amine **47** or its decomposition product, the NH_2 nucleophile, as previously noted (Scheme 2.19). This can be explained based on the steric hindrance, by the adjacent sp3 carbon in **47**, which prevents the nucleophilic attack of the amine **47** to give the product **72**. At the same time, the amine **47** or its decomposition product (at 160 ºC), might serve as base to mediate the reaction of the phenol with the alkene **64**, as the case mentioned in Figure 2.8.

The second transformation that occurred (Scheme 2.22) was the reduction of the nitro group of **64** to give the aniline **72**. No metal catalyst was present in the reaction to catalyse such a reduction, however, the catalyst free reduction of the aromatic nitro group is possible and has been reported previously,[169] where aromatic nitro compounds can be heated with a hydrogen source at 110 ºC for 24 h, in a solvent like water or water/DMSO mixture.[169] Amines and hydrazines are well known sources of hydrogen in reduction reactions,[170,171] and therefore, the amine **47** or its decomposition products are proposed to be the hydrogen source for the reduction of the nitro group in **64** to the amino group in **72**, when heated at 160 ºC for 14 h in a sealed tube. The sealed tube also provided the pressure required for such a reduction reaction. However, because the product **72** is out of the scope in this work, the reaction was not repeated on the other alkene **62** or **63**, to investigate the scope of this reduction.

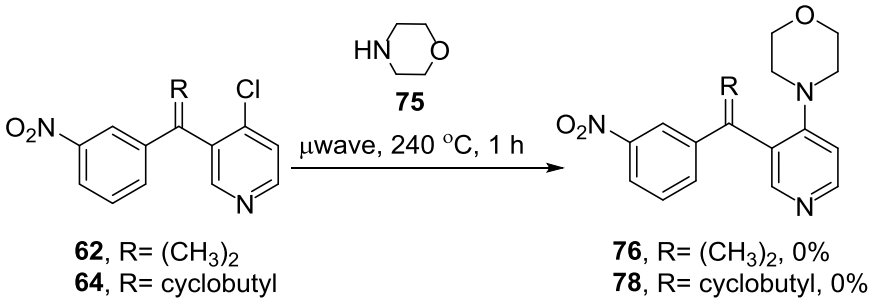
2.5.5. *Amination of the alkenes 62-64 using morpholine 75*

Replacing the amine side chain **33** and the amine **47** by morpholine **75** was an applicable strategy that was previously used.[157] The AAPM **36** carrying the morpholine side chain (Figure 2.2), displayed a weak to moderate activity against the dengue virus (64% inhibition at 25 µM).[157] The synthesis of the AAPM **36** was also achieved in 62% yield.[157] Therefore, the alkenes **62-64** were reacted with neat morpholine **75**, by heating the mixture at reflux for 48 h (Scheme 2.23). Excess morpholine **75** was then removed by distillation under vacuum, leaving oily residues. Only the cyclopropyl analogue **77** was isolated after being subjected to alumina column chromatography. Analysis of the ESI-MS (negative ionization) spectrum of **77** showed a peak at 336.9 *m/z* which corresponded to $[M\text{-}H^+]^-$. Analysis of the ^{13}C NMR spectrum of **77** showed a peak at 66.8 ppm assigned to the morpholine C2 and C6, and two carbon resonating at 53.4 ppm assigned to C3 and C5 of the morpholine ring in **77**.



Scheme 2.23 Amination reaction of the alkenes **62-64** with morpholine **75**.

In case of the alkenes **62** and **64**, no reaction was detected and the alkenes were fully recovered after the reaction time (Scheme 2.23). The ability of the alkene **63** to react with morpholine **75** was related to the less hindrance, in comparison with the more hindered alkenes **62** and **64**. Therefore, harsher conditions were employed on alkenes **62** and **64**. Neat morpholine **75** was heated with either alkene **62** or alkene **64**, in a microwave tube (Scheme 2.24). The maximum heat applied was 240 ºC (using the maximum microwave power), and the reaction was held for 1 h. The reaction was cooled with compressed air for 10 min. TLC and low resolution mass analysis indicated no change in the reactants.



Scheme 2.24 Attempted microwave reaction of the alkenes **62**, **64** with morpholine **75**.

The failure to synthesise **76** and **78** was mainly attributed to the steric hindrance by the dimethyl groups in case of alkene **62**, and the bulky cyclobutyl of the alkene **64**. However, the alkenes **62**, **64** were found to be stable after heating at 240 ºC for 1 h, and were fully recovered.

2.6. Concluding remarks

This chapter presented the attempted trials to synthesise the nitro AAPM derivative **46**, starting from the nitro-isobutyrophenone **38**. The sequence of the reaction sequence proceeded well until the final amination step with the racemic side chain amine **33**. Different conditions were tried for this amination, however, the polymerization of the side chain and the loss of the starting material, the aminochloropyrimidine **43**, provided additional difficulties in repeating the reaction. During the optimization of the dehydration reaction for the alcohol **40**, using conc. H_2SO_4, a new furopyrimidine **42** was obtained in an excellent yield. This is considered a novel short synthetic pathway to access the furopyrimidine skeleton which has been rarely reported in the literature.

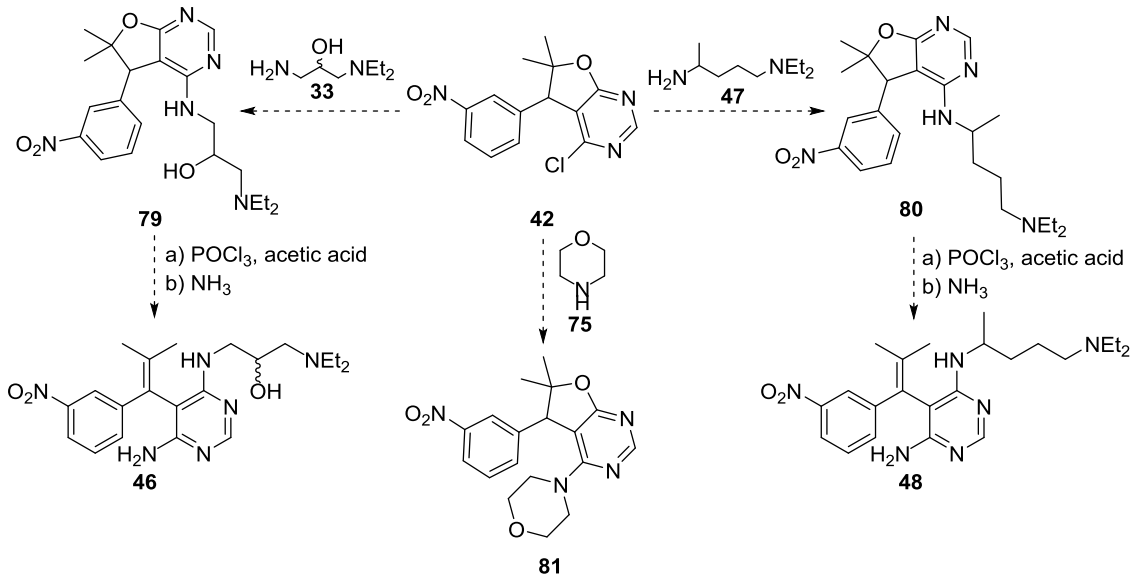
The racemic amine **33** was then replaced by the amine **47**, which contains features present in the anti-malarial chloroquine **8**. Final amination using **47** was not successful due to the instability of **47**, decomposing and releasing NH_2 as a nucleophile, and also due to the steric hindrance caused mainly by the two methyl groups in the aminochloropyrimidine **43**. One way to decrease the steric hindrance was by tethering the two methyl groups by using the less hindered cyclopropyl group. The replacement took place, and the alcohol **50** was isolated, however, it was found to be unstable, and the conversion to the alkene **51** was not successful. Another possibility is to use a less hindered analogue of **47** that does not have the methyl germinal to the NH_2 group.

A modification of the AAPM structure was then achieved by replacing the 4,6-dichloropyrimidine **28** by the simpler 4-chloropyridine HCl **55**. The addition and dehydration reactions proceeded well and three alcohols **59-61** and the corresponding alkenes **62-64** were obtained in good yields. However, the final amination with the amine **47** was not successful with any of the alkenes, including the Pd catalysed amination chemistry. Instead, the ether-aniline product **72** was obtained when trying the amination with the alkene **64**, using phenol as a solvent. This product showed that the nucleophilic substitution on the alkenes **62-64** was possible, although steric hindrance prevented the addition of the amine **47** in the desired position. Product **72** showed also the instability of the amine **47** when placed in harsh conditions which might be required for such nucleophilic substitution reactions, as can be noted by the reduction of the nitro group on alkene **64** to an amino group in **72**, depending on the amine as a hydrogen source.

Finally, the amination with the secondary amine, morpholine **75**, was performed on the alkenes **62**-**64**, however, only the cyclopropyl analogue **77** was isolated in good yield. This confirmed the difficulties in performing such reactions in the presence of hindered systems.

In conclusion, the alkene **63**, with the less hindered cyclopropyl group, represents the most promising alkene to perform more amination reactions under moderate conditions. The synthesized compounds, not only represent milestones in the optimization process of the AAPM derivative syntheses, but also provide new scaffolds for testing against the CHIKV as well as against malaria, to identify new hit compounds.

Some alternative routes to the AAPM derivatives are to be considered. The synthesized new furopyrimidine **42** can be considered for the amination reaction, as it is less stericaly hindered than the corresponding alkene **41** (Scheme 2.25). The furan ring can then be open using $POCl_3$[160] to give the final aminated products.



Scheme 2.25 Alternative synthesis of the AAPM **46**, **48** and **81** using the furopyrimidine **42** intermediate.

CHAPTER 3: Modelling Study of the CHIKV nsP3 as a Possible Drug Development Target

3.1. Introduction

The function of *alphaviruses* nsP3 has remained unknown, although mutations can affect different steps of the viral replication machinery.[75] It is constructed of two domains, with the first being a unique macro domain in the conserved N-terminal region. The C-terminal region is less conserved and is phosphorylated in about 16 positions on serines and threonines.[76,77] The function of phosphorylation is not understood, but it was found that deletion of these phosphorylated residues decreases the RNA synthesis level in infected cells.[78]

The N-terminus of nsP3 contains a macro domain (known also as X domain), which binds to ADP-ribose derivatives and RNA, and is able to hydrolyse ADP-ribose-1"- phosphate,[80,81] a side product of cellular pre-tRNA splicing. Therefore, it is believed to control the metabolism of ADP-ribose 1"-phosphate and/or other ADP-ribose derivatives which have regulatory functions in the cell. The ADP ribose-binding site within the nsP3 macro domain is solvent-exposed and points away from the other domains in the nsP23 polyprotein. Based on sequence conservation in *alphaviruses*, it has been shown that residues just after the nsP3 macro domain play a role in positioning of the nsP23 complex cleavage site.[82] It can be inferred from the crystal structure of the nsP23 precursor protein of the closely related *alphaviruses*, SINV, that the nsP2 is connected to the nsP3 through the macro domain of the nsP3.[83]

In 2010, the crystal structure of the nsP3 macro domain for the CHIKV was solved[80] (Figure 3.1). It is formed of 672 residues and contains six-stranded β sheets (cyan colour in Figure 3.1) and three α helices (red colour in Figure 3.1). The 2D intermolecular interactions between the residues in the binding pocket of the enzyme and the ADP-ribose,[85] as analysed from the crystal structure, are also shown (Figure 3.2).

As shown in Figure 3.2, the key binding residues are: Arg144, Asp10, Ile11 The111, Gly112, Ser110, Tyr114, Val113, Asn24, Asp31 and Val33. The binding complex is formed of 10 H-bonding and one π-cation interaction.[85] The PO_4^{2-} moiety

showed the strongest interactions with these residues in the enzyme pocket. Also, the ribose (with Thr111) and the diphosphate (with Gly112, Ser110, Tyr114, Val113, Val 33) units were found to play major roles in the CHIKV nsP3 ADP-ribose complex.[85]

The nsP3 of CHKV uses a conserved proline-rich motif to interact with the Src-homology-3 (SH3) domain of amphiphysin-1 and amphiphysin-2 proteins, of the host cell, which are related members of the BAR (Bin-Amphiphysin-Rvsp) protein superfamily implicated in several cellular functions.[86]

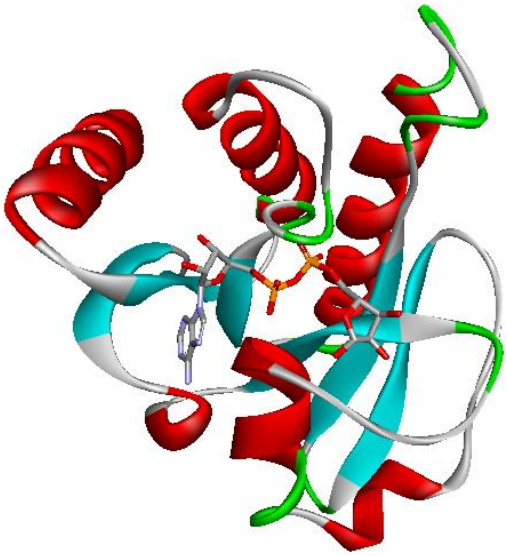


Figure 3.1 Crystal structure of the CHIKV macro domain with the bound ADP-ribose, generated from the PDB file code: 3GPO,[80] β sheets are shown in cyan, while the three α helices are shown in red colour.

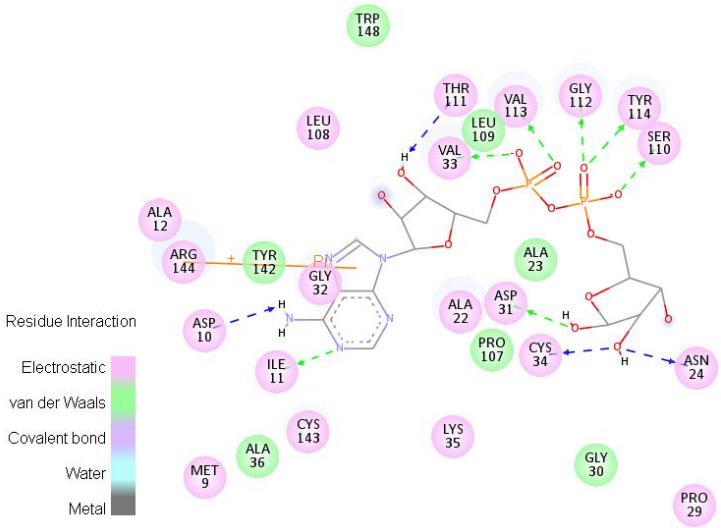


Figure 3.2 2D representation of the interaction of ADP-ribose inside the nsP3 macro domain binding pocket showing the other residues inside the active site (Generated from the PDB file: 3GPO[80]).

It was recently[87] found that the nsP3 acts as an inhibitor of stress granule assembly by recruiting G3BP into cytoplasmic foci. The conserved SH3 domain-binding motif in nsP3 (the proline-rich motif, not included in the crystal structure) is also essential for both nsP3-G3BP interactions and viral RNA replication. G3BP (Ras GTPase-activating protein-binding protein) is an enzyme in human cells and a member of the heterogeneous nuclear RNA-binding proteins.[88] This protein plays a major role during infection and in the assembly of stress granules. Stress granules are membranous cytoplasmic focal structures (foci) that immediately aggregate in response to cellular stress; this last action leads to impaired translation of most mRNAs.[89] These stress granules may have antiviral activity that is inhibited by CHIKV replication by the nsP3 SH3 domain-binding motif.[87]

The crystal structure of the CHIKV nsP3 macro domain represents only one domain of the whole protein, however, the main function of the CHIKV nsP3 is located within this X domain, through the ADP-ribose binding site. Moreover, like other *alphaviruses*, the nsP2 is connected to the nsP3 through the macro domain of the nsP3.[83] Therefore, targeting the binding site of this domain will have effects on both the functions of the whole nsP3 itself, and also, will impair the functions of the polyprotein complex (nsP23 cleavage site). Therefore, impairing the functions of the nsP3 macro domain would either affect the functions of the whole protein (nsP3) as an individual functioning unit, or would impair the functions of the polyprotein complex before being cleaved by the protease activity of the nsP2. In either case, the viral replication machinery will be impaired.

This chapter describes the investigation of using of the CHIKV nsP3 as a potential drug target by utilizing the ADP-ribose binding site to perform virtual screening of a compound library. In this structure-based search approach, the co-crystallized ligand (ADP-ribose) was used as a control. Ligands which showed higher scores within the pocket were considered as potential competitive inhibitors, and ligands which showed lower affinities were discarded.

3.2. Results and discussion:

3.2.1. *The data set library used in the search*

The freely available **NCI diversity set III** was selected for this virtual screening. The diversity set was derived from almost 140,000 compounds available in plates for high

throughput screening approaches. Only available compounds were considered. The 71,756 compounds meeting this criterion were filtered using the program Chem-X (Oxford Molecular Group). Chem-X uses defined centres (hydrogen bond acceptor, hydrogen bond donor, positive charge, aromatic, hydrophobic, acid, base) and defined distance intervals to create a particular finite set of pharmacophores. 3-Point pharmacophores were used resulting in almost 1,000,000 possible pharmacophores. The Chem-X diverse subset generating function reads through a set of structures and determines the acceptable conformations. For each acceptable conformation, it determines all the possible pharmacophores. The pharmacophores for the current structure are compared to the set of all pharmacophores found in structures already accepted into the diverse subset. If the current structure has more than a preset number of new pharmacophores, it is added to the diverse subset. The requirements were set as 5 new pharmacophores and, additionally, 5 or fewer rotatable bonds. This procedure resulted in the selection of 1990 compounds.*

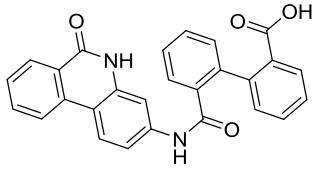
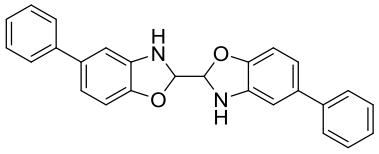
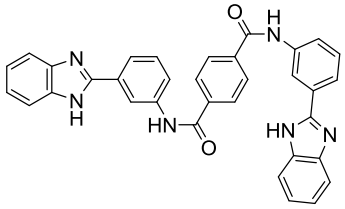
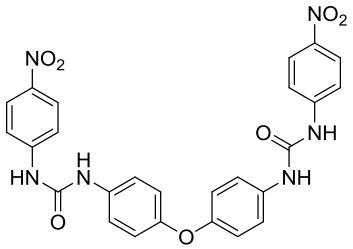
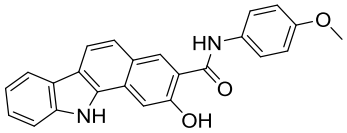
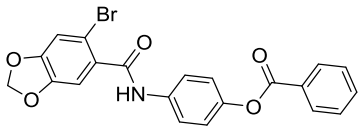
3.2.2. *Virtual screening with the CHIKV using Autodock Vina*

In this chapter, Autodock Vina (Vina)[172] was used to perform virtual screening of the NCI Diversity Set III into the ADP-ribose binding site within the CHIKV nsP3 macro domain crystal structure (pdb: 3GPO[80]). Vina is freely available to the academic community. It is two orders of magnitude faster than Autodock 4.0 (Autodock[173]). Vina performs well relative to Autodock; while Autodock is slightly better at predicting the energy of binding (standard error of 2.2 kcal.mol^{-1} versus 2.8 kcal.mol^{-1} for Vina), Vina more accurately reproduces cocrystallized ligand poses.[172,174] Vina has been used successfully for virtual screening with reliable results.[175]

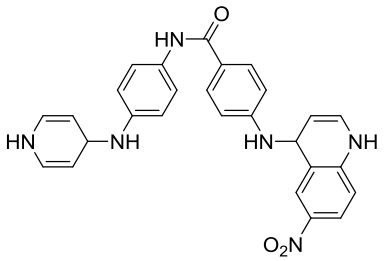
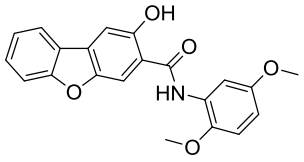
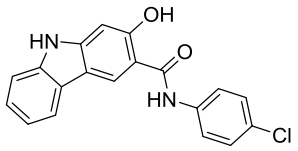
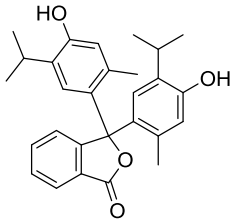
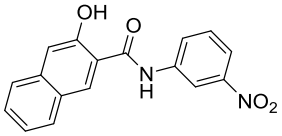
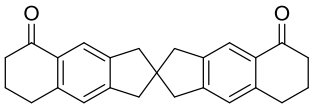
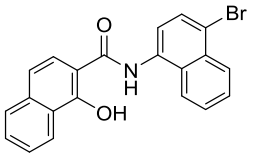
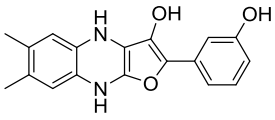
Table 3.1 shows the top 17 ranked docked poses that achieved lower binding energies (ΔG_{bind}) than the co-crystallized ligand (ADP-ribose), with the binding energies (Kcal/mol), interacting residues and calculated log (ClogP). The validation of the docking accuracy was achieved by docking the native co-crystallized ligand (ADP-ribose) into its binding site of nsP3 macro domain. The result was satisfactory where the docked ligand was exactly superimposed on the native co-crystallized ligand with a RMSD value of 0.34 Å, and a binding free energy of -10.4 kcal/mol. The top 10 compounds were selected for *in vitro* antiviral evaluation against the CHIKV.

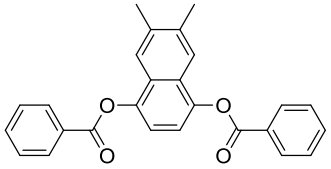
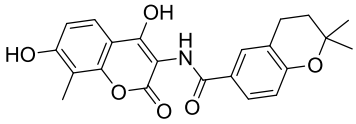
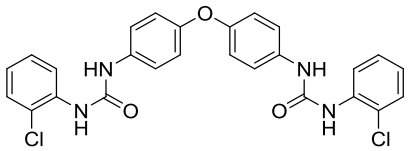
* The **diversity set III** definition and the preparation procedures were extracted from the NCI website.

Table 3.1 Autodock Vina top ranked molecules which scored higher than ADP-ribose, binding affinity (ΔG_{bind}), interacting residues and calculated logP (ClogP).

Entry	Structure	ΔG_{bind} Kcal/mol	Interaction Residues	ClogP*
1		-12.0	Leu108 (H bond, 1.99 Å), Ser110 (H bond, 2.26 Å), Arg144 (cation-π)	2.94
2		-11.9	Trp148 (π-π), Leu108 (H bond, 2.2 Å)	2.41
3		-11.5	Arg144 (H bond, 2.15 Å), Ser115 (H bond, 2.01 Å), Thr111 (H bond, 2.18 Å), Trp148 (π-π)	6.77
4		-11.4	Asn24 (H bond, 2.24 Å), Cys34 (H bond, 2.08 Å), Val113 (H bond, 1.83 Å), Leu108 (H bond, 2.47 Å), Arg144 (H bond, 1.84 Å), Ile11 (H bond, 2.42 Å), Tyr142 (H bond, 2.30 Å)	6.75
5		-11.3	Ile11 (H bond, 2.42 Å)	3.01
6		-11.1	Trp148 (π-π), Val113 (H bond, 2.03 Å), Asp31 (H bond, 2.39 Å)	5.13

* Calculated using ChemBioDraw Ultra 13.0 (2012).

7		-10.9	Arg144 (cation-π, H bond, 2.09 Å), Ile11 (H bond, 2.27 Å), Tyr142 (H bond, 1.88 Å), Thr111 (H bond, 2.19 Å), Gly112 (H bond, 2.02 Å)	3.43
8		-10.8	Thr111 (2 H bond, 1.81, 2.48 Å), Gly112 (H bond, 2.26 Å)	4.72
9		-10.7	Trp148 (π-π), Arg144 (H bond, 1.91 Å), Thr111 (H bond, 1.83 Å), Gly112 (H bond, 2.40 Å)	3.54
10		-10.7	Agr144 (cation-π)	4.51
11		-10.7	Arg144 (H bond, 2.13 Å), Asp145 (H bond, 2.29 Å), Thr111 (H bond, 1.89 Å), Gly112 (H bond, 2.22 Å)	2.09
12		-10.6	Asp31 (H bond, 1.90 Å)	4.91
13		-10.6	Arg144 (cation-π, H bond, 2.48 Å), Thr111 (H bond, 2.09 Å), Gly112 (H bond, 2.17 Å), Val33 (H bond, 2.36 Å)	4.39
14		-10.6	Leu108 (H bond, 1.93 Å), Tyr142 (H bond, 2.15 Å), Thr111 (H bond, 1.79 Å), Gly112 (H bond, 2.33 Å),	3.78

15		-10.6	Leu108 (H bond, 2.44 Å), Val113 (H bond, 2.32 Å), Tyr114 (H bond, 2.33 Å)	6.94
16		-10.6	Val33 (H bond, 2.47 Å), Thr111 (2 H bond, 1.70, 1.88 Å), Gly112 (H bond, 2.24 Å), Arg144 (cation-π)	3.51
17		-10.4	Asn24 (H bond, 2.19 Å), Thr111 (H bond, 2.16 Å), Val113 (H bond, 2.23 Å), Arg144 (cation-π)	6.88

From Table 3.1, it can be seen that all the top ranked poses could interact with the key residues inside the CHIKV nsP3 pocket, which interact with ADP-ribose (Figure 3.2). The results also revealed the participation of new residues within the pocket, where the amino acid Leu108 was found to either accept H-bonds through the carbonyl group, or donate H-bonds through the NH group with poses 1, 2, 4, 14 and 15 (Figure 3.3). The Leu108 residue is located inside the pocket close to the key residues Ser110, Thr111, Gly112 and Val113. Interactions with these residues were found to be important for the binding of the ADP-ribose ligand.[85] Therefore, interactions with Leu108, should also strengthen the interactions within this site of the pocket. This would stabilizes the ligand-enzyme complex at this site, therefore, competing well with ADP-ribose (did not show interactions with Leu108) to fit within the pocket inside the CHIKV nsP3 macro domain.

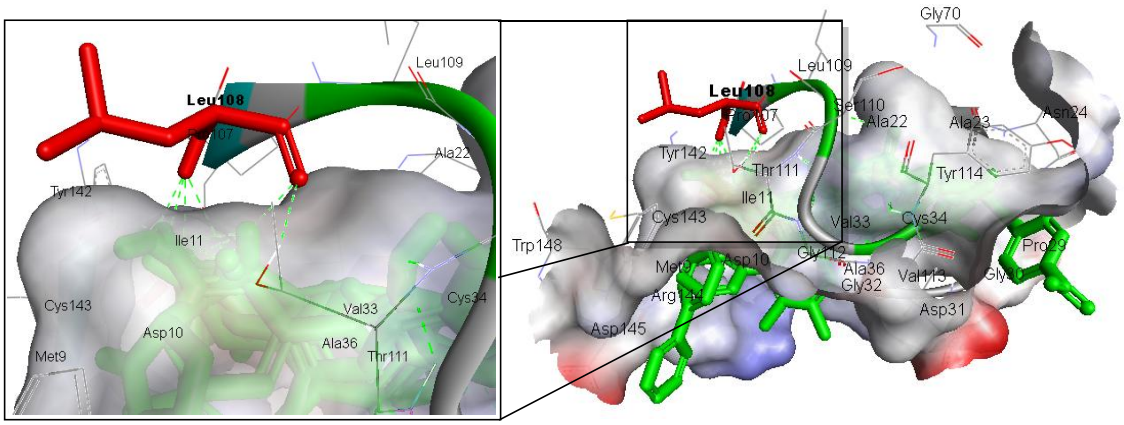


Figure 3.3 Leu108 (red colour) interactions within the nsP3 pocket with entries 1, 2, 4, 14 and 15 (poses are shown in green colour). Leu108 can either accept H-bonds through the backbone CO or donate H through the backbone NH group.

Another common residue, Trp148 which is lying on the pocket floor, was found to have π-π stacking interactions with entries 2, 3, 6 and 9 (Table 3.1). These poses were found to position aromatic systems against the indole ring of Trp148, giving stability of the aromatic residues against the pocket mouth (Figure 3.4). The amino acid residue Asp145 was also found to form a H-bond only with entry 11 (Table 3.1). Ser115 was also found to form a H-bond only with entry 3.

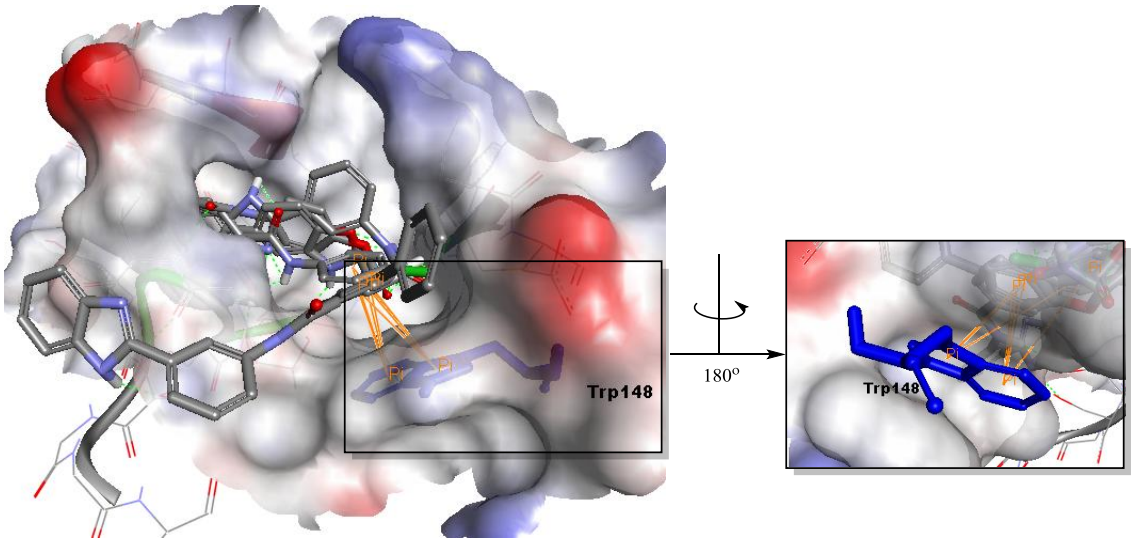
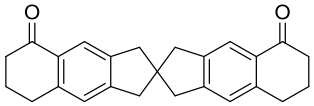
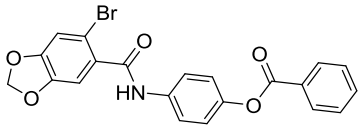
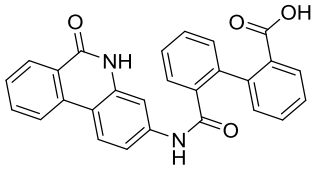
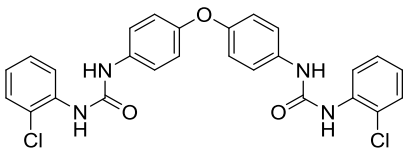


Figure 3.4 Trp148 π-π interactions with the aromatic rings of entries 2, 3, 6 and 9. Trp148 is shown in blue colour within the transparent surface of the nsP3 pocket. The π-π interactions are shown as orange lines.

3.2.3. *Ranking of the Autodock Vina output using Autodock*

To select the hit to be optimized, the hit list obtained by Autodock Vina (Table 3.1) was subjected to ranking using the Autodock algorithm, which was then used for the *in silico* optimization. Re-docking of the hits using Autodock was performed to re-score the list with the Autodock scoring function, using the native ligand (ADP-ribose) as a control, and excluding those hits which are ranked below the score of ADP-ribose. Table 3.2 shows the re-docking results. Autodock 4 results for the ADP-ribose: binding energy (ΔG_{bind} = -11.07 Kcal/mol) and predicted inhibition constant (K_i = 7.74 nM).

Table 3.2 Hits which passed the Autodock re-ranking, achieving scores higher than (ADP-ribose), their binding energies (ΔG_{bind}), predicted inhibitory constant (K_i) and interacting residues. Entries of the table are the same as in Table 3.1.

Entry	Structure	ΔG_{bind} *Kcal/mol*	Predicted K_i (nM)	Interaction residues
12		-12.66	0.528	Asp31 (H bond, 1.74 Å)
6		-12.01	1.56	Asn24 (H bond, 2.32 Å), Asp31 (H bond, 2.15 Å), Leu108 (H bond, 1.90 Å), Val113 (H bond, 2.01 Å), Tyr142 (π-π)
1		-11.71	2.59	Leu108 (H bond, 1.97 Å), Thr111 (H bond, 2.21 Å), Arg144 (cation-π)
17		-11.37	4.62	Ala22 (H bond, 2.09 Å), Cys34 (H bond, 1.71 Å), Gly112 (H bond, 2.16 Å), Val113 (sigma-π), Arg144 (H bond, 2.06 Å)

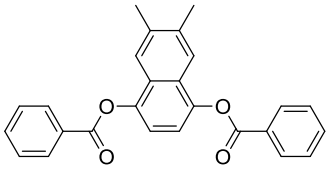
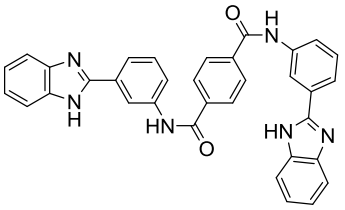
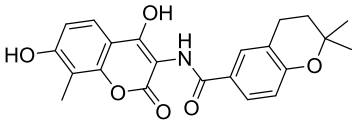
15		-11.35	4.75	Leu108 (H bond, 2.09 Å), Val113 (H bond, 2.16 Å), Tyr114 (H bond, 2.10 Å), Tyr142 (π-π)
3		-11.31	5.1	Ala22 (H bond, 2.36 Å), Gly112 (H bond, 2.48 Å) Val113 (H bond, 2.11 Å), Arg144 (2 cation-π, H bond 1.97 Å)
16		-11.14	6.81	Asp31 (H bond, 2.38 Å), Val33 (H bond, 1.78 Å), Leu108 (H bond, 1.74 Å), Thr111 (H bond, 2.27 Å), Gly112 (H bond, 2.09 Å), Val113 (H bond, 2.18 Å), Tyr114 (H bond, 2.01 Å), Agr144 (H bond, 2.23 Å)

Table 3.2 showed that the amino acid residue Asp31 is critical for the ligand-enzyme interaction complex as demonstrated by entries 12 and 6, where both ere predicted to form H-bonding with Asp31 with predicted distances of 1.74 Å and 2.15 Å, respectively. Entry 12 also achieved shape matching and therefore was predicted to have the highest ΔG_{bind} value. Entry 16 showed the ability to form a H-bonding with Asp31 with a predicted distance of 2.38 Å, a greater distance from those observed in entries 12 and 6. Although ranked the last pose (Table 3.2), entry 16 showed 8 H-bonds with key binding residues within the pocket, ranging from 1.78 to 2.38 Å.

Entry 3 (benzimidazole derivative) was selected for the docking based optimization due to its easy chemical access, high scoring in the first Autodock Vina screening, highly optimizable structure and it showed the ability to form 2 cation-π interactions with Arg144 (similar to ADP-ribose). It was also selected in order to validate the efficacy of the optimization process (optimizing a lower ranked hit rather than a higher ranked one).

3.2.4. *In silico optimization of entry 3*

Hit 3 (Table 3.2) showed interaction within the CHIKV nsP3 pocket (ΔG_{bind} = -11.31 Kcal/mol, predicted K_i = 5.1 nM). Figure 3.5 shows the 2D representations of the interaction of entry 3 with the nsP3 residues. This hit can be described as two benzimidazol-2-yl phenyl units connected to a central phenyl ring through amide linkages. One of the benzimidazol-2-yl phenyl forms important interactions with the key amino acid residue Arg144, where the benzimidazole moiety forms the π-cation stacking interaction with the cationic side chain of Arg144 *via* the benzo moiety, and another π-cation interaction with the phenyl ring, whereas one N atom of the imidazole moiety accepts a H-bond from one of the Arg144 NH_2 (Figure 3.5). The other benzimidazole moiety can donate a H-bond, through the imidazole moiety to the backbone carbonyl group of Ala22. One of the amide carbonyl (Figure 3.5) accepts two H-bonds from the backbone amide NH of the key residues Gly112 and Val113.

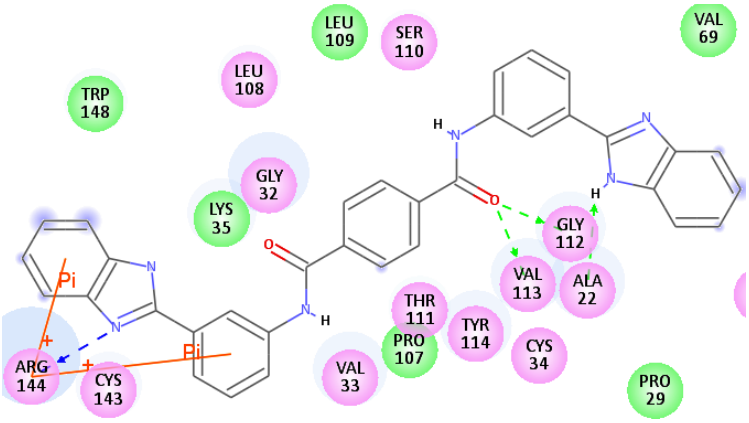


Figure 3.5 2D representations of the interaction of entry 3 within the CHKV nsP3.

Therefore, the important moieties for hit 3 (Table 3.2) appear to be the benzimidazol-2-yl phenyl moieties and the amide linkages. The ClogP for this hit is 6.77, which is due to the presence of 5 phenyl rings in the structure. Removal of the central phenyl ring was the first modification performed on hit 3, producing pose_3_1 (Figure 3.6) which suggested more positive interactions with active site residues (ΔG_{bind} = -11.44 Kcal/mol, predicted K_i = 4.08 nM), a slight improvement in the Autodock score compared to hit 3. ClogP for pose_3_1 is 4.97, lower than that of the original hit 3.

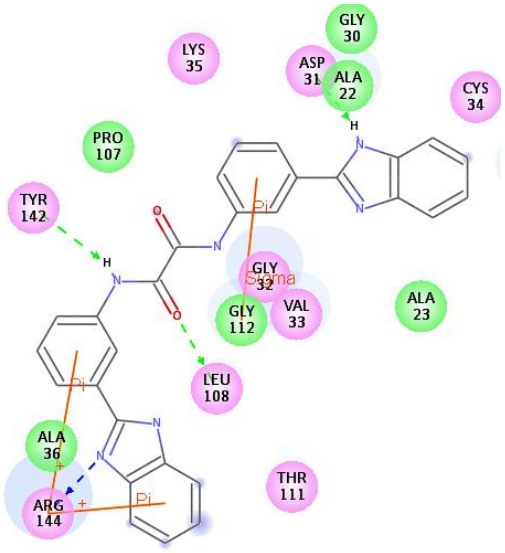


Figure 3.6 2D representations of the pose_3_1 within the CHIKV nsP3.

Figure 3.6 shows that pose_3_1 has the same predicted interactions with the key residue Arg144. However, Gly112 also interacted, not through a H-bond as in case of hit 3 (Figure 3.5), but through a sigma-π stacking interaction (Figure 3.6). Two amino acid residues participated in new H-bonds, the key residue Tyr142 (1.82 Å) and the backbone NH of Leu108 (1.88 Å). The other terminal benzimidazole N did not donate a H-bond to Ala22 as in hit 3, instead, a H-bonding interaction with the key residue Asp31 (Figure 3.6) was predicted. The 3D model of pose_3_1 suggested further possibilities of H-bonds with the polar hydrogens of the key amino acid residue Asn24 (Figure 3.7), through one of the benzimidazole rings. Such H-bond formation will require the presence of an acceptor on the terminal benzo moiety of the benzimdazole close to Asn24 (Figure 3.7). Moreover, the presence of an acceptor on the other benzimidazole moiety (which forms interactions with Arg144), would make possible a H-bond with the polar hydrogens of Lys35 (Figure 3.7).

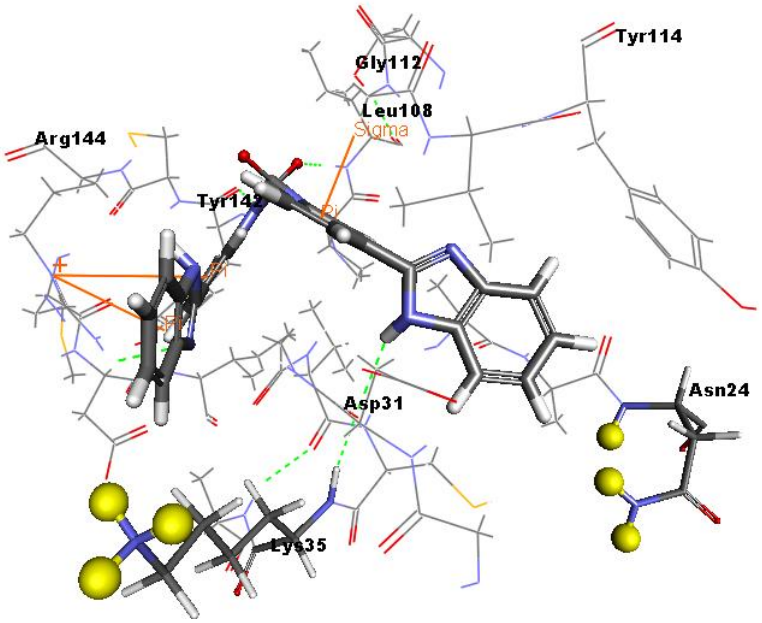


Figure 3.7 3D representation of pose_3_1 within the nsP3 pocket. Asn24 and Lys35 polar hydrogens are shown as yellow spheres.

Therefore, the next modification was to insert an acceptor group on the terminal benzimidazole rings. This acceptor was to be added on both benzimidazole rings, to achieve the two possible interactions with Asn24 and Lys35. Several acceptor groups such as NH_2, OH, COOH and NO_2 could be added on the desired position, however, the nitrile group (CN) was selected as an acceptor to be added on the benzimidazole ring. It was selected because it is able to form relatively strong H-bonds with the receptor residues,[176] and also, the presence of the carbon atom would act as a spacer, bringing the N closer to the receptor residues.[177] Therefore, pose_3_1 was optimized to give Pose_3_2. This modification, not only represented a better *in silico* binding profile (ΔG_{bind} = -13.12 Kcal/mol, predicted K_i = 239.69 pM), but also, it showed interactions with 9 key residues within the nsP3 binding pocket (Figure 3.8). Furthermore, the ClogP for pose_3_2 was brought down to be 3.84 compared to 6.77 of the original hit 3.

a) b)

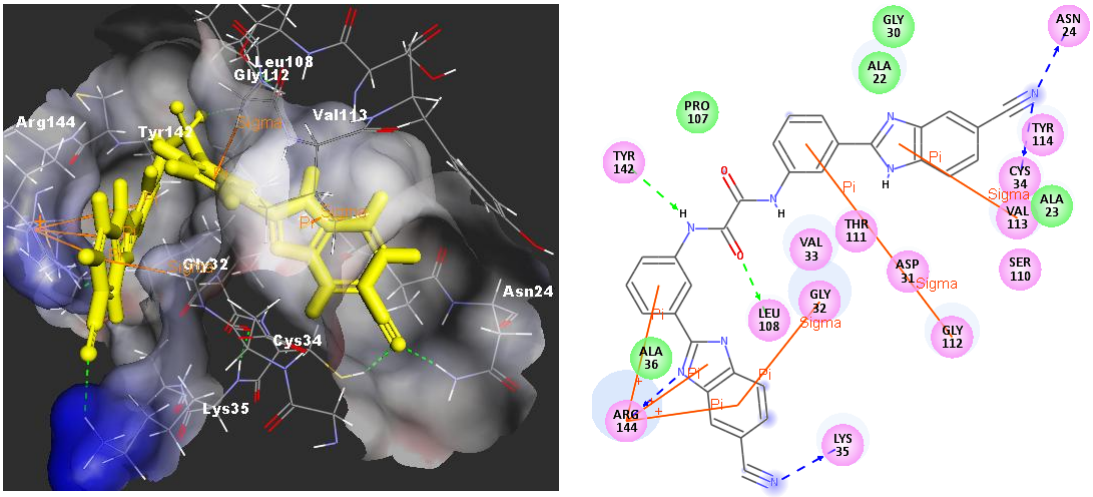


Figure 3.8 Pose_3_2 within the CHIKV nsP3 pocket. a) 3D representation, b) 2D representation. H-bonds are shown as green dashes, cation-π and sigma-π interactions are shown as orange lines.

As predicted, the nitrile group introduced on the benzimidazole rings, not only accepted a H-bond from Asn24 (1.90 Å), but also accepted another H-bond from the key residue Cys34 (2.19 Å). The nitrile group on the other benzimidazole ring also accepted a H-bond from Lys35 (2.35 Å) as predicted. It is envisaged that such interactions with Cys34 and Lys35 residues would stabilize the pose within the pocket cavity openings (Figure 3.8a). The pocket central residues Gly112, and Val113 were able to form sigma-π stacking interactions with a phenyl ring, and the imidazole moiety, respectively (Figure 3.8b). The molecular weight of pose_3_2 is 522.5, with ClogP value of 3.84, and these drug-like properties made it an attractive candidate for synthesis and testing against the CHIKV.

3.3. Synthesis of pose_3_2 (86)

The retrosynthetic analysis of pose_3_2 (Figure 3.9) suggested that it could be accessed through three synthetic steps; the synthesis of the nitro-benzimidazole, which can be then reduced to the amino derivative, and reacted in a final step with oxalyl chloride (Figure 3.9) to give the bis-product.

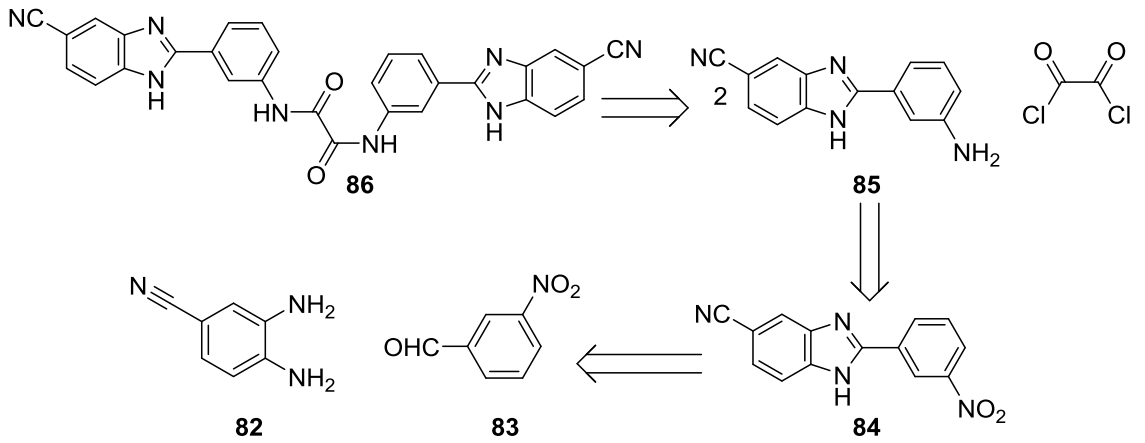
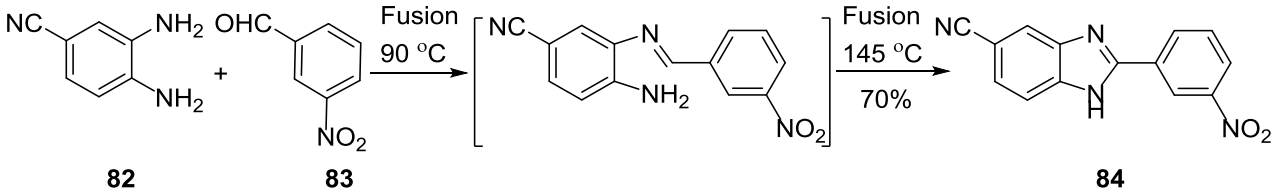


Figure 3.9 The retrosynthetic analysis for the synthesis of pose_3_2 (**86**).

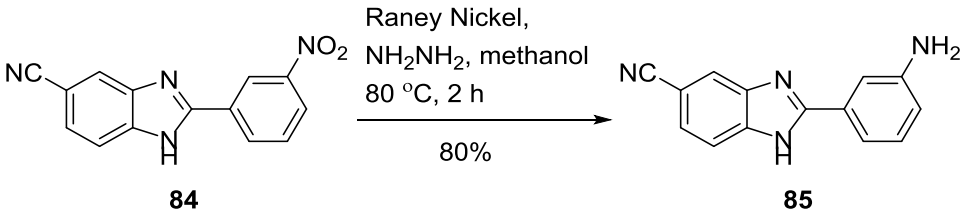
The synthesis of the benzimidazole nucleus can be achieved by a variety of methods. The synthesis can involve the reaction between the phenylene diamine derivatives with aldehydes,[178] carboxylic acids[179] or acid chlorides,[180] utilizing a variety of different solvents and catalysts. In this work, the benzimidazole **84** was synthesized *via* the reaction of the 3,4-diaminobenzonitrile **82** with 3-nitrobenzaldehyde **83** in solvent free conditions, where the two solids were mixed together (1:1 molar ratio) and the mixture was heated with vigorous stirring at 90 ºC. Both solids melted, giving a dark brown oil mixture (Scheme 3.1). After a few minutes of stirring the dark oil at 90 ºC, the mixture solidified as a dark brown solid. At this stage, the first step of the reaction occurred which was the condensation reaction of the aldehyde **83** with one of the amines in **82** giving the Schiff base intermediate (Scheme 3.1). Increasing the temperature to 145 ºC initiated the cyclization step, where the 2nd amine attacked the electron deficient azomethene carbon, where the dark brown solid condensation intermediate melted again giving a dark red oily mixture. Stirring continued for about 3 min, after which the mixture solidified again as a dark red solid (Scheme 3.1). The solid was then extracted into boiling ethanol, and crystallized from a mixture of methanol, 1,4-dioxane and water to afford the benzimidazole **84** as a yellow crystalline solid.



Scheme 3.1 The synthesis of the benzimidazole **84**.

Analysis of the ESI-MS spectrum of **84** showed a peak at 264.1 *m/z* that corresponded to (M+). The ^{1}H NMR spectrum showed a broad singlet peak resonating at 13.8 ppm which was assigned for the NH. This solvent free conditions enabled the access to the desired benzimidazole in a short time (less than one hour) and relatively good yield (70%).

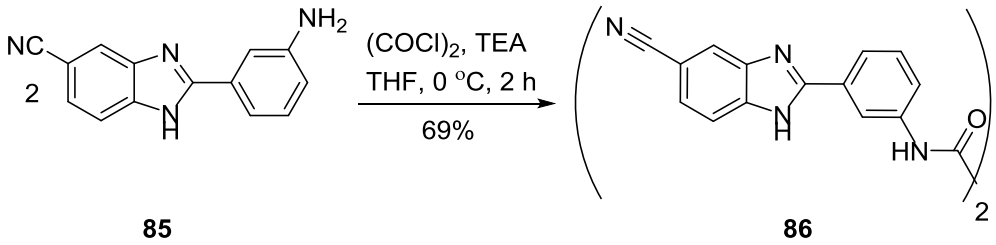
The next synthetic step was the reduction of the nitro group in **84**. Therefore, the benzimidazole **84** was suspended in methanol, and Raney Nickel catalyst was added under N_2 atmosphere with vigorous stirring. The hydrogen source for the reduction was hydrazine monohydrate in methanol, which was added dropwise. During the addition of hydrazine hydrate, hydrogen gas was evolved and upon cessation, the reaction was heated to 80 ºC for 2 h (Scheme 3.2). The mixture was then filtered through celite while hot, and methanol and excess hydrazine hydrate were removed under reduced pressure to give the amine **85** as a dark yellow solid.



Scheme 3.2 Reduction of the benzimidazole **84** to the amino derivative **85**.

The analysis of ^{1}H NMR spectrum of the product **85** showed a broad peak resonating at 5.37 ppm which was assigned as the NH_2 group. ESI-MS (negative ionization) showed a peak at 233.0 which was assigned for (M-1).

The final synthetic step was the dimerization of the amine **81** using oxalyl chloride. The amine **85** was dissolved in dry THF in the presence of TEA, and the mixture was stirred under N_2 atmosphere at 0 ºC. A solution of oxalyl chloride in dry THF was then added dropwise with vigorous stirring. After addition, the mixture was left to warm to room temperature with stirring for 2 h (Scheme 3.3) after which the mixture became turbid. The reaction was then quenched with chilled water and a solid precipitated immediately. The solid was filtered and washed with water and methanol, recrystallised from DMF to afford the final pose_3_2 (**86**) in 69% yield. The ^{13}C NMR spectrum of the product showed a peak at 158.8 ppm which was assigned to the CO groups. The ^{1}H NMR spectrum showed a peak at 11.08 ppm which was assigned for the two oxalamide NH groups. The IR spectrum also showed a band at 2219 cm^{-1} that was assigned for the nitrile group.



85 **86**

Scheme 3.3 Reaction of the amino benzimdazole **85** with oxalyl chloride to afford the final compound **86**.

In this reaction (Scheme 3.3), only the primary amine NH_2 group was involved in the reaction with oxalyl chloride. The benzimidazole NH is reactive and might react with acid chlorides, however, it is known that *N*-acylbenzimidazoles undergo hydrolysis to the free NH again in the presence OH groups or water.[181] The rate of this hydrolysis increases if the imidazole or benzimidazole carries electron withdrawing groups such as the nitro group.[182] The electron withdrawing nitrile group of **85** is therefore believed to achieve similar effect, deactivating the acylation of the benzimidazole NH.

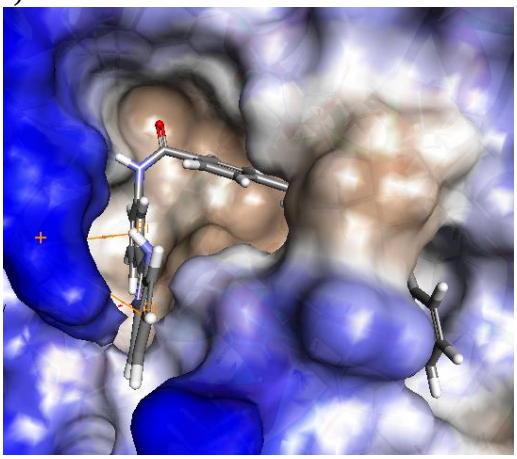
3.4. Concluding remarks

This chapter discussed the use of the macro domain (domain X) of the CHIKV nsP3 as a possible target for the drug design against that virus. The nsP3 is a multifunctional viral protein, with the macro domain being the domain which regulates the ADP-ribose

substrates within the cellular process. Furthermore, the ADP-ribose binding site within the macro domain is close to the connection site with the nsP2. Therefore, occupying the ADP-ribose binding pocket with an exogenous ligand (inhibitor) should affect the function of the nsP3 either in a polyprotein complex or as an individual functioning protein. In this chapter, a virtual screening process of the NCI diversity set III, a library that carries drug-like molecules, was performed using the ADP-ribose pocket as the target binding site. Autodock Vina, a fast and reliable virtual screening engine, was used and 17 hits were identified as potentially stronger enzyme binders compared to the original ligand (ADP-ribose) which may serve as inhibitors for the CHIKV nsP3 functions. The search revealed that the hits made interactions with the identified key residues within the pocket, and also with new residues that tended to stabilize the ligand-receptor complex, such as Leu108, Trp148, Ser115 and Val22. The top 17 hits (Table 3.1) were obtained from the NCI and are currently under antiviral evaluation (at Griffith University).

The top hits were then subjected to re-ranking against the original ligand (ADP-ribose) using Autodock 4. This re-ranking resulted in 7 hits achieving higher scores than the ADP-ribose. One hit, hit 3 (Table 3.2) was selected for further docking based optimization. This optimization resulted in pose_3_2 (**86**) which showed a significantly improved *in silico* profile compared to the original hit (Figure 3.10).

a) Hit 3

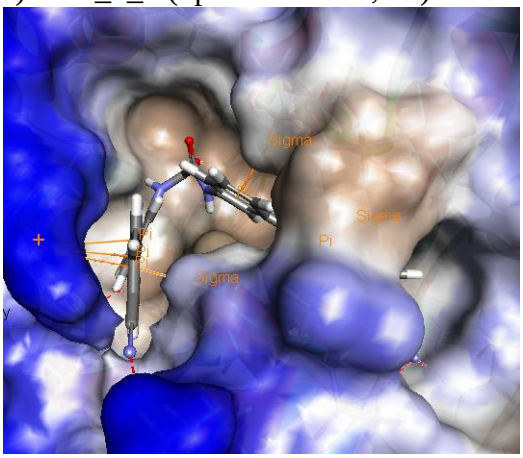


ΔG_{bind} = -11.31 Kcal/mol

K_i = 5.1 nM

ClogP = 6.77

b) Pose_3_2 (optimized hit 3, **86**)



ΔG_{bind} = -13.12 Kcal/mol

K_i = 0.239 nM

ClogP = 3.84

Figure 3.10 Comparison between (a) hit 3 and (b) the optimized structure **86**.

The optimized structure **86** was then accessed through three facile synthetic steps. This optimized hit **86** represents an early stage towards the discovery of a lead compound, based on the structure-based approach, which can be further optimized using *in silico* and *in vitro* in-depth studies.

The virtual screening hits 4 and 17 (Tables 3.1, 3.2) have the same linear skeleton and showed good interactions with the CHIKV nsP3 ADP-ribose pocket. If these molecules were demonstrated to be promising in activity, different derivatives can be accessed through one step reaction (Figure 3.11), and therefore, they might be a subject of both *in silico* and *in vitro* optimizations.

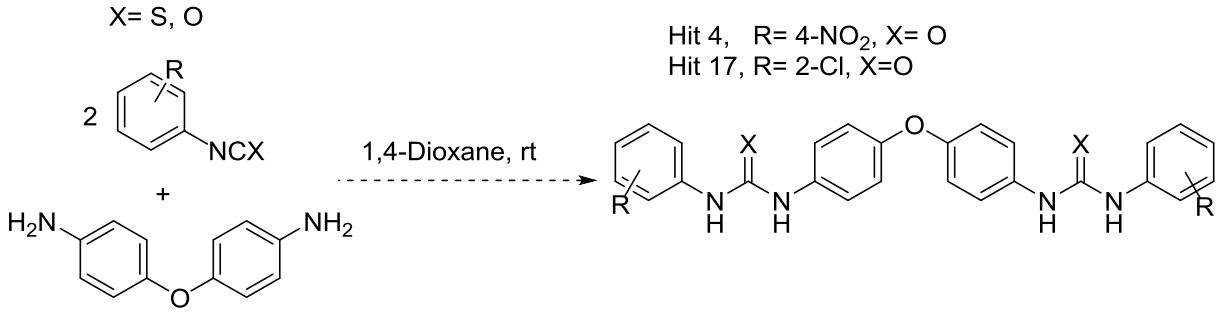


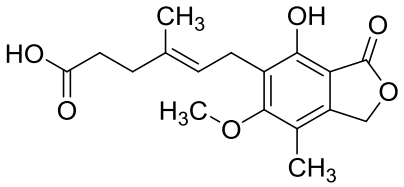
Figure 3.11 Possible access of different analogues for hits 4 and 17.

The inhibitor discovery status for the CHIKV nsP3 is still in the very early stages. Thus fare, the enzyme is not even available as a testing kit to validate the potential inhibitors. Efforts should be devoted to develop such validation techniques. Protein crystallization with the inhibitor would enhance the drug design approach for this virus.

CHAPTER 4: Synthesis of Mycophenolic Acid Analogues as Inhibitors of the Chikungunya Virus

4.1. Introduction

Mycophenolic acid (**15**, Figure 4.1) was isolated about one hundred years ago from the cell culture of *Penicillium brevi-compactum*.[125] It acts as an inhibitor for the inosine monophosphate dehydrogenase (IMPDH) enzyme which is involved in the *de novo* biosynthesis of guanine nucleotide. It has good anti-proliferation activity and has been established as an anticancer[126] and antiviral agent, as well as an immunosuppressant properties.[127]



15

IC_{50}= 0.2 μM against CHIKV

Figure 4.1 Chemical structures of mycophenolic acid **15**.

Recently, MPA was shown to inhibit the CHIKV replication and virus induced cell death with an IC_{50} value of 0.2 μM and a selectivity index value of 150 (compared to vero cell lines),[128] and it was found to induce CHIKV apoptosis. It works by inhibiting IMPDH, and when the CHIKV infected cells were provided with exogenous guanosine (GMP), MPA was prevented from inducing the CHIKV cell death.[128] Considered as a good lead compound, the *in vivo* activity of MPA acid against CHIKV should be investigated as it is well known that mycophenolic acid suffers from a metabolic drawback associated with rapid conjugation of the C-7 phenolic hydroxyl group with glucuronic acid.[183,184,129,130]

MPA as a lead: Mycophenolic acid **15** has good drug like properties with a molecular weight of 320.3, calculated logP (clogP[*]) value of 2.92, calculated LogD[†] value of 3.56

[*] Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)

[†] Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA, LogD is a pH dependent distribution coefficient version of the logP. It represents the drug permeability (lipophilicity) at a certain

(pH=1-2), 0.67 (pH=7.4) and 2.52 (pH=5.5) indicating good blood solubility and oral bioavailability (oral drug absorption occurs in the small intestine at pH ~ 5.5[185,186]). With the good CHIKV activity (IC_{50} = 0.2 µM), MPA **15** is a promising template for developing anti-CHIKV drugs, with the metabolic drawbacks to be avoided for better bioavailability properties.[183,184,129,130] The structure of CHIKV IMPDH enzyme is not yet available and therefore the structure based drug design for inhibitors is not a possibility. Instead, a ligand based drug design approach can be used starting with MPA **15** as a lead compound.

4.2. Design of MPA analogues

Mycophenolic acid consists of three major components (Figure 4.2), an aromatic head (the 3-oxo-1,3-dihydroisobenzofuran moiety), a rigid linker (the alkene connection) and an acidic tail (the carboxylic acid group). The aromatic head carries two major hydrogen bond acceptors (the carbonyl group and the oxygen atom of the lactone ring). The aromatic benzo moiety carries a hydroxyl group, a methoxy group and a methyl group.

In this study, the three components of MPA **15** were investigated as optimizable components that could be changed with a variety of similar systems, starting from simple cores such as benzoic acids and anilines. Figure 4.2 shows the possible access to MPA **15** analogues starting from such derivatives. Both the acids and anilines could be converted into the aromatic head groups carrying H-bond acceptor capacity, and then could be connected through simple C-C coupling, to the other two cores, the linker and the acidic tail. R groups on the acids and anilines were selected to give a similar environment as in **15**.

physiological pH. Oral drugs to be able to be absorbed by passive diffusion through the gut wall, should have their lipophilicity within a given range between 1 and 4 on LogD scale. Note that LogD = LogP at neutral pH, therefore LogD describes the actual LogP values at different pHs.

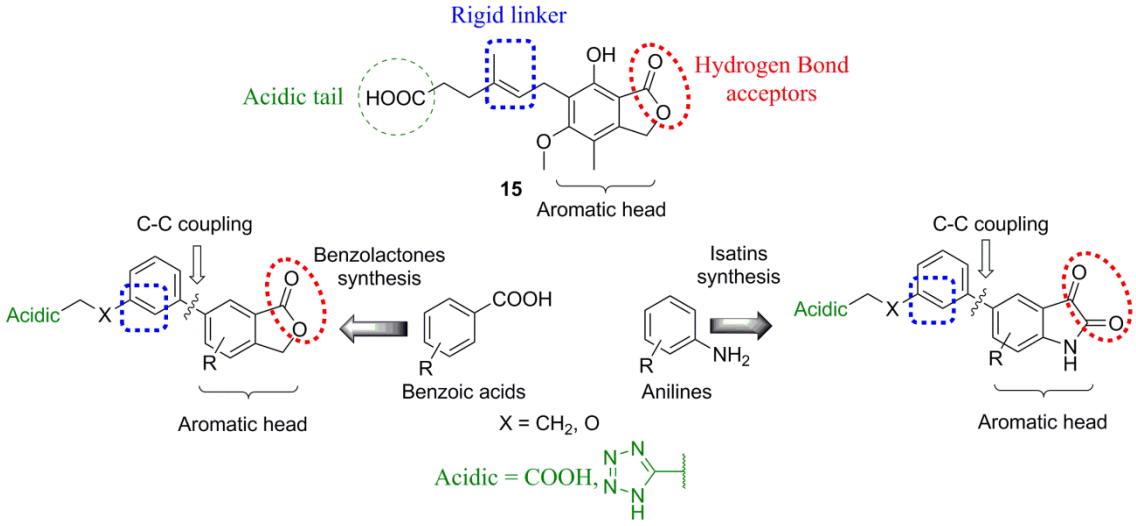


Figure 4.2 Design of MPA **15** analogues starting from benzoic acids and aniline derivatives. The red circles indicate the HBA groups, the blue boxes indicate the rigid linker and the green circle refers to the acidic tail moiety which could be either COOH or the bioisostere tetrazole.

4.2.1. Design Strategy

The design of alternative and similar cores to MPA **15** was based on:

1. Changing the aromatic head: the aromatic head was either kept as a benzolactone or changed to an isatin moiety (Figure 4.2). The use of benzolactones as the aromatic head groups would highlight the importance of this functionality for the anti-CHIKV activity. Isatins were also selected as alternative possible replacements for the aromatic head group; as they have been showed to possess various pharmacological and biological properties including anticholinesterase, anticonvulsant, antiinflammatory, antihypertensive, antihypoxic, antimicrobial, and antiviral properties.[187,188] Furthermore, the isatin moiety has been found in nature, e.g. as a component of the secretion from the parotid gland of *Bufo* frogs,[187] and in humans as a metabolic derivative of adrenaline.[189] Isatin is also found in the brain.[190] The structure of isatin participates in the activity, where the two carbonyl groups (red circles, Figure 4.2) could accept H-bonds from residues in the target receptors, and the NH can donate a H-bond to acceptor residues.[191,192] The aromatic core π-system of isatin was also found to participate through aromatic interactions with some target receptor residues.[193] Therefore, replacing

the benzolactone head of MPA **15** with isatin, would have some structural similarities, with possible similar interactions with the yet unknown viral receptor.

2. Replacing the rigid linker alkene with the rigid phenyl ring (blue boxes, Figure 4.2) would be an applicable strategy for obtaining a rigid linker between the aromatic head and the acidic tail of MPA analogues. Utilizing the C-C cross coupling chemistry[194,195,196] could facilitate such facile access to the biphenyl system. This phenyl aromatic linker may also provide a π-electron system for interaction within the yet unknown target site.

3. The acidic tail was kept as either a carboxylic group or changed to the bioisosteric tetrazole[197] moiety (green, Figure 4.2). The carboxylic group is highly polar and can ionize and hinder absorption. The tetrazole ring has similar physiochemical properties but offers some advantages over carboxylic acids. Like COOH groups, tetrazole rings have an ionisable acidic proton at pH 7.4. They are also planar in structure, however, the tetrazole anion is 10 times[198] more lipophilic than the carboxylate anion and this can enhance drug absorption as a result. Tetrazole rings are also resistant to many of the metabolic reactions that occur to the carboxylic groups.[199,198]

This chapter discusses the different synthetic strategies undertaken for the access to the aromatic head groups (benzolactones and isatins) and the C-C coupling reactions utilized to tether the synthesized head groups to the linker and acidic tails (either COOH groups or tetrazoles) producing different conjugates. It also discusses the calculated physicochemical properties of the synthesized target compounds as templates for developing drug-like candidates.

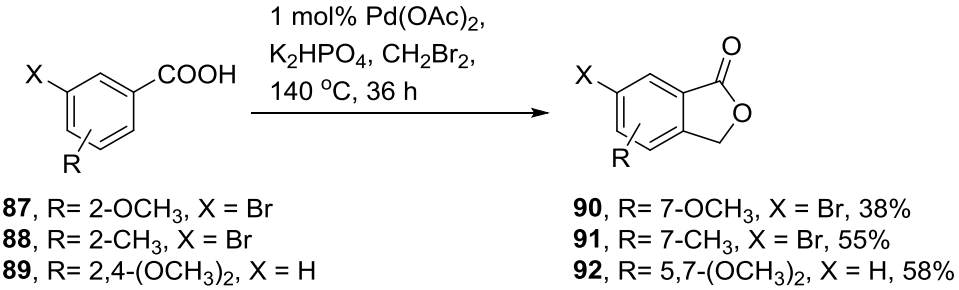
4.3. Synthesis of the Aromatic Heads

4.3.1. Synthesis of the isobenzofuran-1(3*H*)-one (benzolactones)

Many procedures have been reported for the synthesis of the isobenzofuran-1(3*H*)-one ring system including from 2-formylbenzoic acid,[200] phthalaldehyde,[201] 1,2-phenylenedimethanol[202] and from benzoic acids.[203]

In this study, benzoic acids were used as starting materials.[203] Therefore, a mixture of benzoic acid derivatives **87**-**89**, palladium acetate, dipotassium phosphate and dibromomethane as the solvent and carbon donor was heated in a sealed tube at 140

ºC under a N_2 atmosphere for 36 h (Scheme 4.1). After dilution of the reaction with CH_2CH_2, filtration through celite, washing with acid and evaporation of the solvent, the lactones **90**-**62** were obtained in 38-58% yield. The obtained compounds were easily purified through either flash column chromatography or recrystallization from a mixture of chloroform/hexane mixture (1:3, v/v). Analysis of the ^{1}H NMR spectra of the isobenzofuran-1(3*H*)-ones **90**-**92** showed a new singlet peak resonating at 5.17-5.23 ppm that was assigned to the newly formed lactone CH_2 group. Analysis of the ^{13}C NMR spectra showed one carbon resonating at 68.2-86.7 ppm that was assigned to the newly added methylene carbon.



Scheme 4.1 The synthesis of the isobenzofuran-1(3*H*)-one heads **90**-**92**.

The reaction proceeds *via* a Pd(II)-catalysed C–H activation where the coordination of the cation with the carboxylate group forces the Pd(II) centre to chelate in the proximity of the C-H bond *ortho* to the acid substituent (Figure 4.3), a geometry that facilitates C-H cleavage.[203] The alkylation attacks the less hindered *ortho* position to the carboxylic acid group.

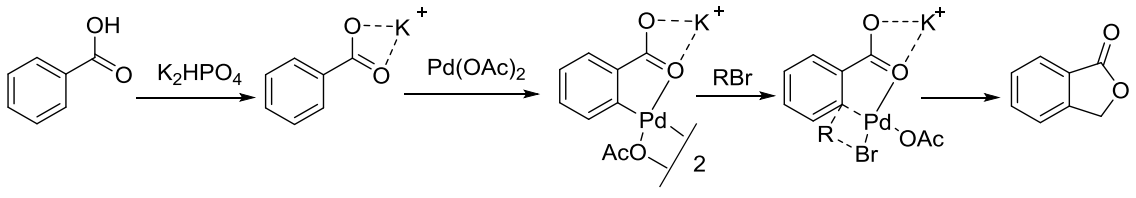
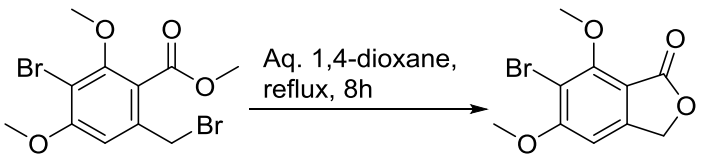


Figure 4.3 Mechanism of benzolactone formation from benzoic acid.[203]

This synthetic procedure was applied either on the brominated benzoic acids, which provided the 6-bromobenzolactones with the coupling handles already intact, e.g.

the 6-bromo-7-methoxyisobenzofuran-1(3*H*)-one **90** and the 6-bromo-7-methylisobenzofuran-1(3*H*)-one **91**, or was applied on the benzoic acid derivative without the bromo substituent, which required further attachment of a coupling handle for the next step, e.g. 5,7-dimethoxyisobenzofuran-1(3*H*)-one derivative **92**.

Addition of coupling handle (bromine) to 92: The presence of the coupling handle (bromine) on C6 of the 5,7-dimethoxyisobenzofuran-1(3*H*)-one **92** was previously described using a different benzolactone synthesis procedure, starting from either 4,6-dibromo-5,7-dimethoxyisobenzofuran-1(3H)-one[204] or from methyl 3-bromo-6-(bromomethyl)-2,4-dimethoxybenzoate (**93**, Figure 4.4).[205] In both cases, the benzolactone skeleton was synthesized with the bromine already attached in the desired position, prior to the cyclization step.

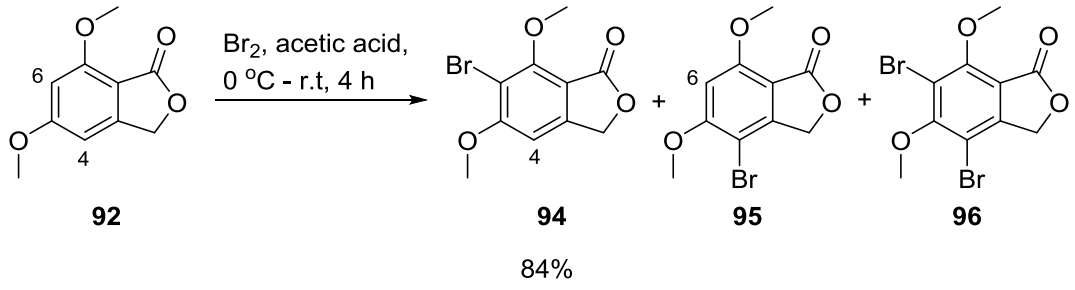


93 **94**

Figure 4.4 Example of the synthesis of the 6-bromo-5,7-dimethoxyisobenzofuran-1(3*H*)-one **94** from methyl 3-bromo-6-(bromomethyl)-2,4-dimethoxybenzoate **93**, with the bromine atom already intact prior to cyclization.[205]

However, the regioselective bromination on C6 of the unsubstituted phthalide **97** (benzolactone) has been previously described in acidic medium, using a mixture of sulphuric acid and trifluoroacetic acid.[206] The selective iodination at C6 was also previously reported.[207] Position 6 in 5,7-dimethoxyisobenzofuran-1(3*H*)-one in **92** is more activated than that of the unsubstituted phthalide **97**, however, it is more sterically hindered by the methoxy groups. In this work, due to the crowding of C6 by the two methoxy groups, a different method was used to access the 6-bromo derivative for this benzolactone **94** from the corresponding benzolactone **92**. A solution of 5,7-dimethoxyisobenzofuran-1(3*H*)-one **92** in glacial acetic acid was cooled in an ice bath before a solution of bromine in glacial acetic acid was added dropwise and the mixture was then stirred at room temperature for 4 h (Scheme 4.2). The resulting solid was filtered and subjected to flash column chromatography (60% CH_2Cl_2 in hexane) to

afford the benzofuran **94**, which was recrystallised from 1,4-dioxane. The isobenzofuran-1(3*H*)-one **94** was obtained in 84% yield with a melting point of 190 ºC (Lit.[205] m.p 206 ºC). The ^{1}H NMR spectrum of **94** showed the disappearance of the C6 aromatic proton which resonated at 6.42 ppm in ^{1}H NMR spectrum of the isobenzofuran-1(3*H*)-one **92**. The EI-MS spectrum of **94** showed also a peak at 272 *m/z* which was assigned to the molecular ion (M+, ^{79}Br). The addition of the bromine solution was carried out at 0 ºC to minimize the production of the 4-bromo isomer **95** and the 4,6-dibromo derivative **96** side products. These side products were not separated due to overlapping bands on the TLC plates, and low mass recovery of this mixture. The EI-MS spectrum of the mixture showed peaks at 272 *m/z* (^{79}Br) that was assigned to the 4-bromo isomer **95** and at 352 *m/z* (^{79}Br, ^{81}Br) that was assigned to the 4,6-dibromo derivative **96**. The reported melting point for the 4,6-dibromo derivative **96** is 136 ºC[204], and that for the 4-bromo isomer **95** is 245 ºC[205].



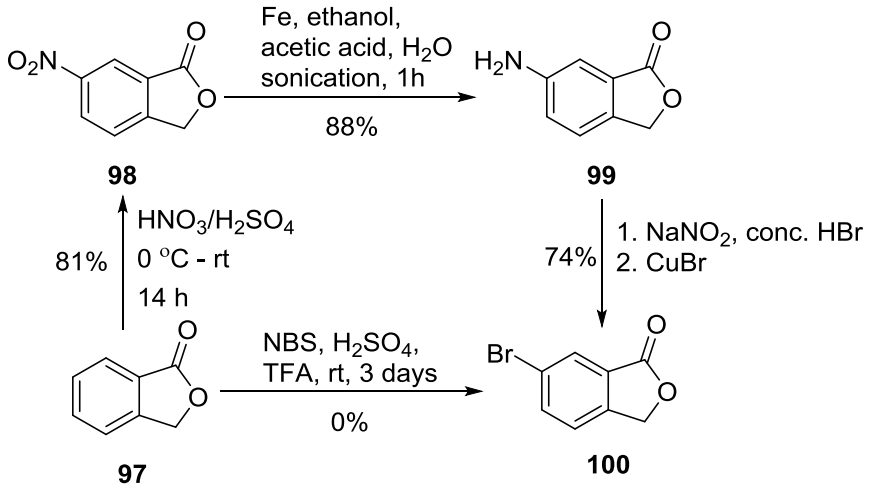
Scheme 4.2 The synthesis of 6-bromo-5,7-dimethoxyisobenzofuran-1(3*H*)-one **90**.

The 2D ^{1}H NMR NOESY experiment of **94** showed a correlation between the aromatic proton H4 with a methoxy group. The methoxy groups appear as singlets resonating at 4.02 ppm and integrate for 6 protons*. No correlation was seen between the aromatic H4 proton and the CH_2 in the NOESY experiment. However, the big difference in the melting points of **94** and **95** argued the bromination of C6.

Access to the unsubstituted benzofuran 100: The synthesis of 6-bromoisobenzofuran-1(3*H*)-one **100** was performed following the reported short procedure[206] initially using *N*-bromosuccinimide (NBS) in the H_2SO_4/TFA mixture and stirring for 3 days (Scheme 4.3). However, this procedure gave a mixture of the multi-brominated phthalide rather

* Different NMR solvents were used to get better resolution of the methoxy groups, however, they always appeared as a singlet peak.

than the 6-bromo derivative **100**. Therefore, in an alternative procedure[208] (Scheme 4.3), phthalide **97** was treated with a nitrating mixture of H_2SO_4/HNO_3 to afford 6-nitroisobenzofuran-1(3*H*)-one **98** in 81% yield. Subsequent reduction with iron[209] in a mixture of ethanol, water and acetic acid, followed by standard diazonium formation and then treatment with CuBr gave the 6-bromoisobenzofuran-1(3*H*)-one **100** in 74% yield,[210] which was recrystallised from methanol.

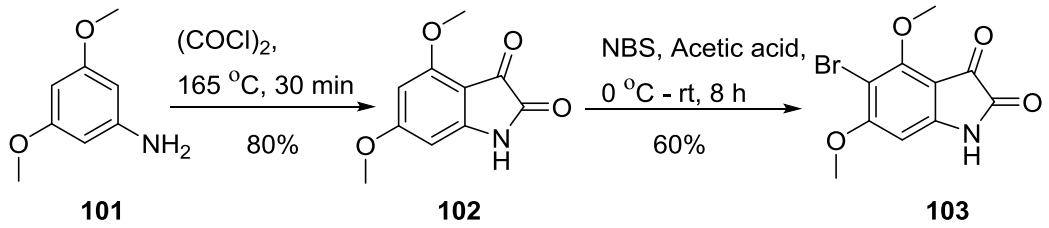


Scheme 4.3 The synthesis of 6-bromoisobenzofuran-1(3*H*)-one **100**.

4.3.2. Synthesis of the isatin aromatic heads

The desired isatin heads were prepared from the corresponding aniline derivatives with or without the bromine coupling handles.

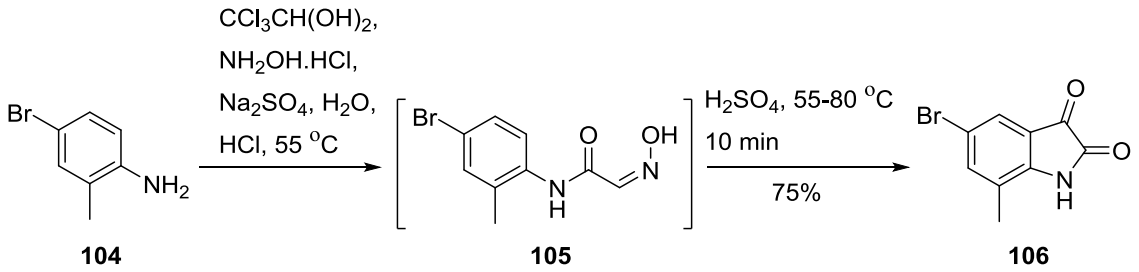
Access to the bromodimethoxy isatin 103: The synthesis of 4,6-dimethoxyindoline-2,3-dione **103** (Scheme 4.4) was undertaken using the Stolle procedure.[211] Therefore, the aniline **101** was reacted with oxalyl chloride to form an intermediate chlorooxalylanilide with spontaneous cyclization, in the absence of Lewis acid, to yield dimethoxyisatins **102** in 80% yield.[212]



Scheme 4.4 The synthesis of 5-bromo-4,6-dimethoxyindoline-2,3-dione **103**.

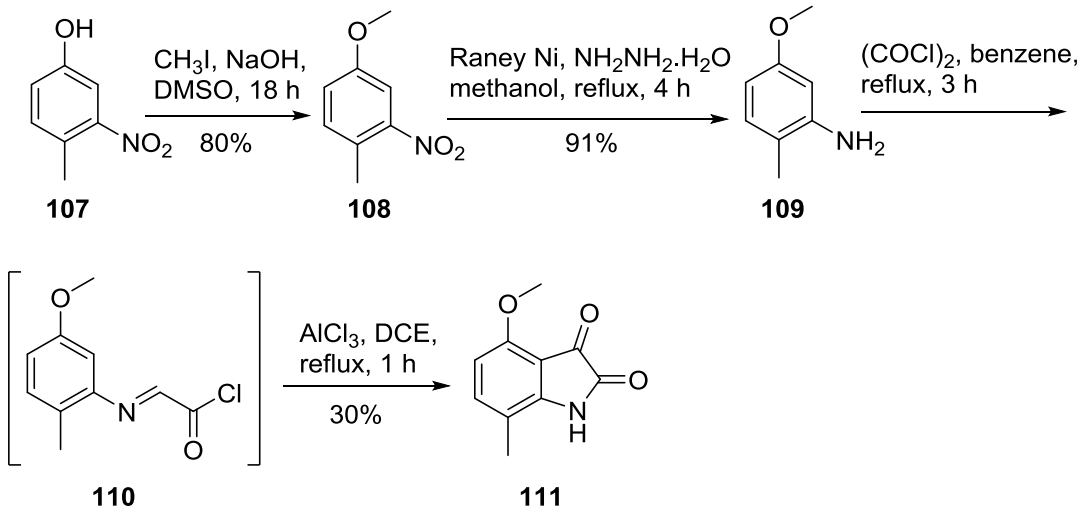
Mono selective bromination of isatins at C5 can be achieved on small scale reactions by using *N*-bromoacetamide in acetic acid medium.[213] Therefore, a solution of the isatin **98** in glacial acetic acid was cooled in an ice bath followed by portion wise addition of the NBS at 0 ºC, and the mixture was then left to warm to room temperature with stirring for 8 h. The yellow solid was filtered, washed with saturated $NaHCO_3$ solution, and then with a water/ethanol mixture and then vacuum dried. The residue was then recrystallised from 1,4-dioxane to afford the bromoisatin **103** in 60% yield. The EI-MS spectrum of **103** showed a peak at 287 *m/z* that was assigned to $(M,^{81}Br)^+$ ion. The ^1H NMR spectrum showed only one aromatic proton peak resonating at 6.36 ppm assigned to H7. This proton correlated to only one methoxy group in the 2D NMR (HMBC and NOESY) experiment. This is only possible when the proton is on C7. The proton in position 5 must show correlation to both methoxy groups which was not observed in **103**. This confirmed that the bromination occurred at position 5 of the isatin ring.

Access to the 7-methylisatin 106: The synthesis of 5-bromo-7-methylindoline-2,3-dione **106** was achieved using Sandmeyer[214] methodology, starting from 4-bromo-2-methylaniline **104** which was reacted with chloral hydrate and hydroxylamine hydrochloride in aqueous sodium sulphate to form an isonitrosoacetanilide **105** (Scheme 4.5), which presented as a gummy solid that was difficult to isolate. This gummy solid was dried and taken directly to the next cyclization step using concentrated sulphuric acid. The mixture was heated gradually to 80 ºC for 10 min. After cooling to room temperature, the mixture was poured onto crushed ice and stirred for 1 h, whereby a solid precipitated and was filtered. The desired compound was isolated in 75% yield over 2 steps. The synthesis of the isatin **106** was previously reported by a different method,[215] where the isatin core was synthesized first, and then was brominated using bromine solution.[215]



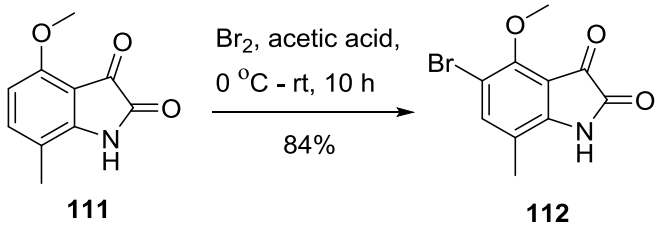
Scheme 4.5 The synthesis of 5-bromo-7-methylindoline-2,3-dione **106**.

Access to methoxymethylisatin 111: The structure of MPA **15** (Figure 4.1) has one methyl group in the *para* position to the OH group on the aromatic head. The isatin 4-methoxy-7-methylindoline-2,3-dione **111** was selected as a similar derivative with the OH group masked as a methoxy functionality, to avoid the metabolic drawback of MPA **15**. The synthesis of this derivative started from the commercially available 4-methyl-3-nitrophenol **107** which was methylated using methyl iodide[216] in DMSO with NaOH as a base to give 4-methoxy-1-methyl-2-nitrobenzene **108** in 80% yield (Scheme 4.6). The 5-methoxy-2-methylaniline[217] **109** was then obtained by reducing the nitro group of **108** over Raney Nickel and using hydrazine monohydrate as hydrogen source, and heating in methanol, to afford the aniline **109** in 91% yield. Reaction of the aniline **109** with oxalyl chloride in benzene gave the intermediate chlorooxalylanilide **110**, without spontaneous cyclization as in the case of the dimethoxy derivative **102** (Scheme 4.4). This was due to the lower reactivity of the monomethoxy intermediate **110**. Therefore, the intermediate **110** was directly suspended in dichloroethane (DCE) and aluminium chloride was then added as a Lewis acid to promote the cyclization step. Acidic workup, followed by extraction with ethyl acetate and concentration of the solvent gave an orange residue that was subjected to column chromatography and elution with 30% ethyl acetate in petroleum spirit afforded the isatin **111** in 30% yield. The IR spectrum of the product showed strong amide absorption peak at 1635 cm^{-1} assigned to the CO group. The EI-MS spectrum showed a peak at 191 assigned to the molecular ion mass (M+). Analysis of the ^{13}C NMR spectrum of **111** showed two additional carbons peaks resonating at 180.9 and 160.1 ppm assigned to the two carbonyls C2 and C3, respectively.



Scheme 4.6 The synthesis of 4-methoxy-7-methylindoline-2,3-dione **111**.

Addition of the coupling handle (bromine) on **111** for the next synthetic step was easily achieved as there was only one dominant bromination site at C5. Therefore, 4-methoxy-7-methylindoline-2,3-dione **111** was reacted with bromine in glacial acetic acid (Scheme 4.7) at 0 ºC to room temperature for 10 h. The resulting solid was collected, washed with $NaHCO_3$, water and then recrystallised from acetic acid to yield the desired isatin **112** in 84% yield. Analysis of the ^1H NMR spectrum of the product showed the disappearance of the H5 that resonated at 6.61 ppm, and the EI-MS spectrum of **112** showed a peak at 271 that was assigned to the $(M, {}^{81}Br)^+$ ion.



Scheme 4.7 The synthesis of 5-bromo-4-methoxy-7-methylindoline-2,3-dione **112**.

4.4. Coupling of the Aromatic Head groups

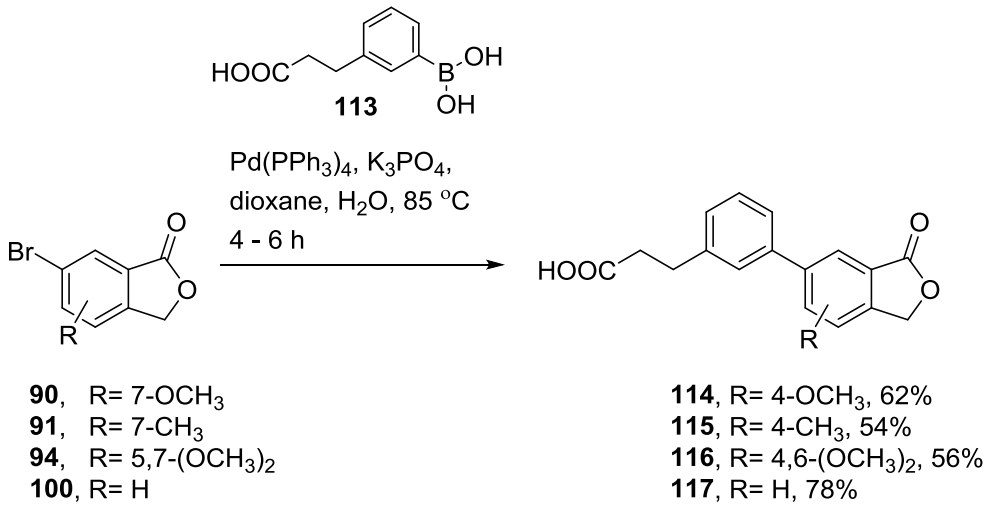
4.4.1. Coupling with the isobenzofuran-1(3*H*)-one

4.4.1.1. Coupling to the linker and COOH tail

Linking of the aromatic head to the acidic tail was accessed utilizing Suzuki C-C cross coupling chemistry,[194,195,196] to access the biphenyl rigid system (Figure 4.2). This facile one step synthesis gave the three main components of the MPA designed analogues: the aromatic head, the rigid linker (phenyl) and the acidic tail (COOH). The chosen boronic acid was the commercially available 3-(3-boronophenyl)propanoic acid **113**. Suzuki cross coupling for the 6-halo-isobenzofuran-1(3*H*)-one derivatives was previously described[206] using a diphasic medium of toluene and water as a solvent, disodium carbonate as a base and tetrakis(triphenylphosphine)palladium(0) as palladium catalyst and heating at reflux temperature.

In this work, 1,2-dimethoxyethane (DME) was initially used as a solvent because of its lower boiling point (~85 ºC) compared to that of toluene (~110 ºC) to retain the stability of the lactone ring when heating with water at refluxing temperatures. Therefore, the 6-bromo-isobenzofuran-1(3*H*)-ones **90**, **91**, **94** and **100** were reacted with 3-(3-boronophenyl)propanoic **113**, potassium phosphate and tetrakis(triphenylphosphine)palladium(0) in DME/H_2O mixture and the reaction was heated in a sealed system at 85 ºC for 8-14 h. After the specified reaction time, the solvent was concentrated and the obtained solids were suspended in HCl and were then extracted with ethyl acetate. After evaporating the solvent, the resulting residues were subjected to PLC (10% methanol in CH_2Cl_2, 0.5% acetic acid) to afford the acid conjugates **114-117**. The reaction worked well, however the isolated yields of the final product were modest (30-35%). This was due to the poor solubility of the benzolactones (**90**, **91**, **94**, **100**) in DME. Interestingly, replacing the DME with 1,4-dioxane and keeping the temperature at 85 ºC (Scheme 4.8), greatly improved the yields up to 78%. Table 4.1 shows the optimized reaction conditions. The IR spectra of the obtained benzolactone-acid conjugates **114-117** showed broad OH absorption bands at 2972-2989 cm^{-1} accompnied by C=O absorption bands at 1699-1755 cm^{-1} indicating the presence of the COOH functional group. Analysis of the ^1H NMR spectrum of 3-(3-(4-methoxy-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenyl)propanoic acid **114** showed a peak at 10.54 ppm that was assigned to the COOH group, and the two propanoic CH_2

groups appeared as two triplets resonating at 2.99 and 2.68 ppm. The purity of the compounds **114-117** was found to be >95% as analysed by normal-phase HPLC using isopropanol with 0.1% TFA in hexane as the developing solvent. The calculated physiochemical properties of this series are shown in Table 4.2.



Scheme 4.8 Coupling procedures for the isobenzofuran-1(3*H*)-one **90**, **91**, **94** and **100** with the boronic acid **113**.

Table 4.1 Optimization conditions of the isobenzofuran-1(3*H*)-ones **114-117** coupling reaction with the boronic acid **113**.

Solvent	Reaction time (h)	Yields (%)
DME/H_2O	8-14	30-35
Dioxane/H_2O	4-6	54-78

The Suzuki coupling is a robust technique for the synthesis of biaryl bonds commonly utilising an aryl boronic acid and an aryl halide. The reaction mechanism of the Suzuki coupling (Figure 4.5) proceeds in 3 key steps; oxidative addition of the Pd(0) complex to the aryl bromide to give the Pd(II) complex, transmetallation of the aryl boronate complex to the Pd(II) complex, and reductive elimination to give the biaryl and restore the Pd(0) catalyst.

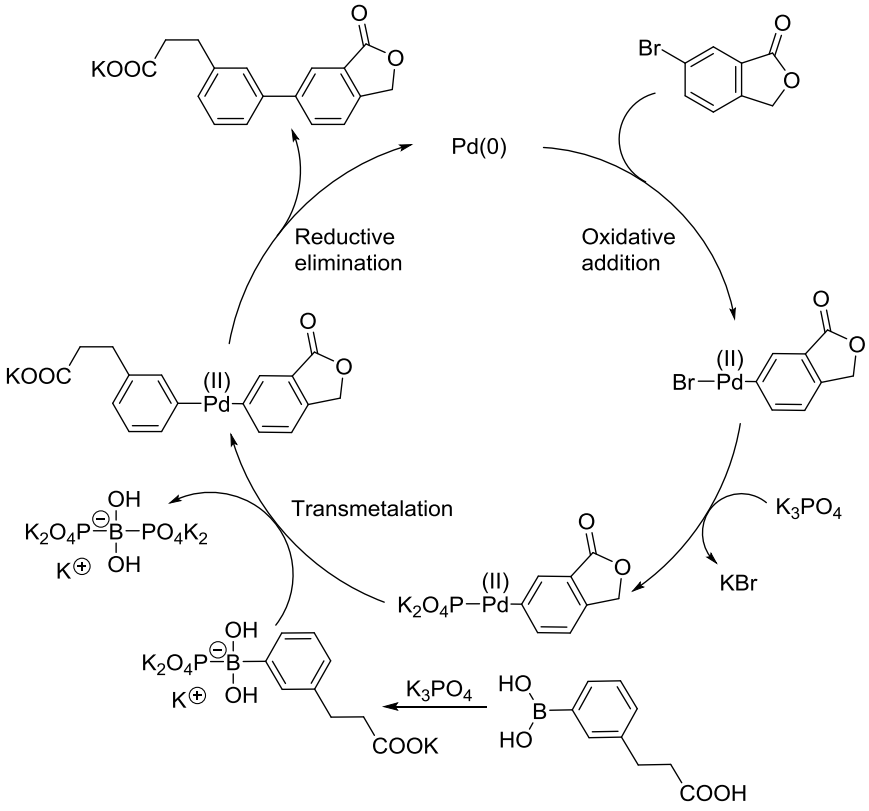
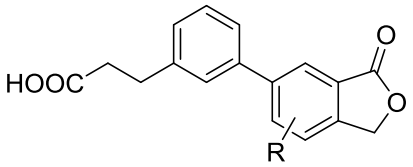


Figure 4.5 The mechanism of the Suzuki reaction to produce the biaryls.

Table 4.2 Physiochemical properties of the benzolactones-carboxylic acid conjugates **114**-**117**.



R=4-CH_3, 4-OCH_3, 4,6$(OCH_3)_2$, H

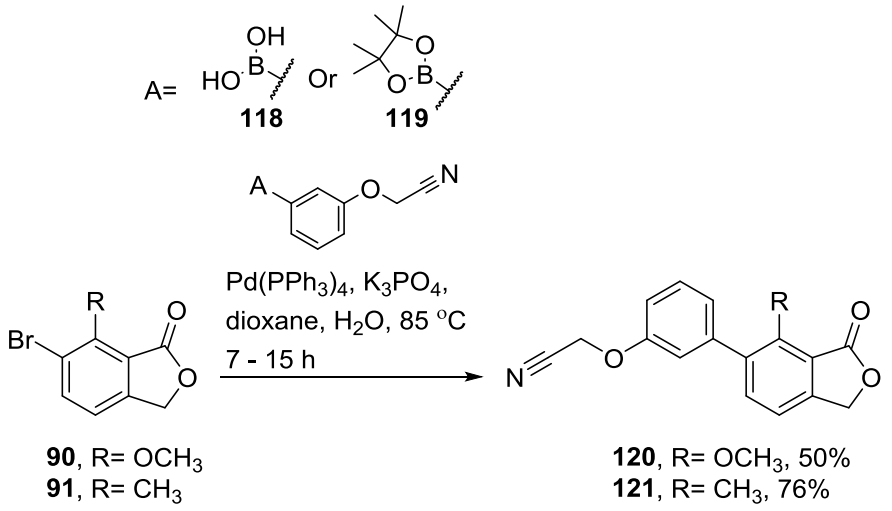
Compound	R	Mwt	clogP*	LogD† at pH		
				1	5.5	7.4
114	4-OCH_3	312.32	1.96±0.59	2.98	1.97	0.14
115	4-CH_3	296.32	2.83±0.52	2.85	2.08	0.34
116	4,6-$(OCH_3)_2$	342.35	1.67±0.60	2.82	1.84	0.01
117	H	282.29	2.37±0.52	2.89	1.90	0.07

* Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)
† Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA

The introduction of the rigid linker (the phenyl ring) was proposed to increase the lipophilicity of the compounds compared to the lead mycophenolic acid which has a LogD value of 0.67 at blood pH. Table 4.2 shows that all the synthesized benzolactone-carboxylic conjugates **114-117** have drug like qualities. This series can ionize in the different pH values as can be seen from the change in the LogD values. In acidic pH (stomach), they show similar LogD values around 2.85 as the ionization will be minimum (unionized), and therefore, **114-117** could be in the absorbable form (can cross the lipid layer). At the blood pH (7.4), the acid conjugates showed low LogD values indicating good blood solubility, and therefore, may present siutable properties for absorbtion from the stomach and distribution through the body with the circulating blood. At the intestine pH 5.5, where the oral bioavailable drugs are absorbed,[186] the ionization shifts towards the lipophilic character without returning to the full unionized state, where the synthesized compounds have good lipophilicity scale (1.84-2.08, lower than that of the lead MPA **15** at the same pH), and therefore, the designed compounds might show good oral bioavailability properties.

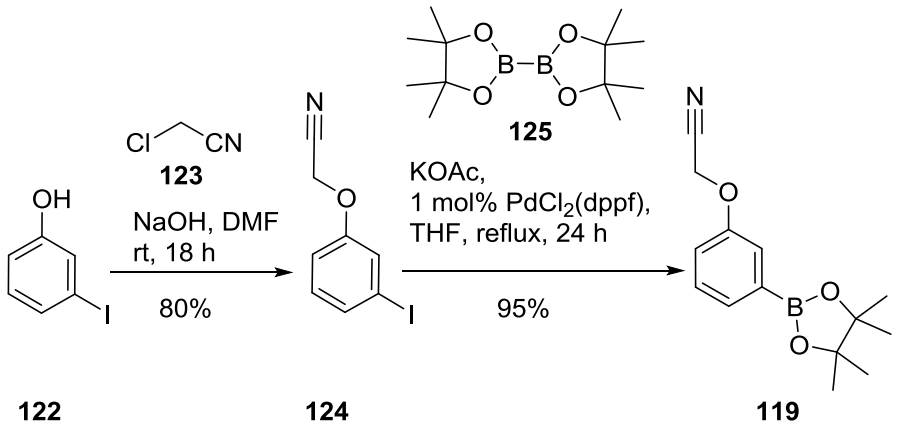
4.4.1.2. Coupling to the linker and tetrazole tail

The first synthesis of the tetrazole ring used the acetonitrile intermediates which can then be cyclised to tetrazoles. Therefore, the commercially available boronic acid, 3-(3-boronophenyl)propanoic acid **118** was utilized in the coupling reaction with the bromobenzolactones. Due to the high cost of **118**, the ester 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetonitrile **119** was also synthesized as an alternative to **118** and was used in the coupling reaction (Scheme 4.9). The same coupling conditions were used: 1,4-dioxane/H_2O mixture, potassium phosphate as a base and tetrakis(triphenylphosphine)palladium(0) as a catalyst, and the mixture was heated under a N_2 atmosphere for 7-15 h. The solvent was then evaporated and the resulting residues were suspended in HCl, and extracted with ethyl acetate. After evaporating the solvent and subjecting the residues to flash column chromatography, the acetonitrile intermediates **120** and **121** were obtained in 50% and 76%, respectively. Analysis of the ^{1}H NMR spectra of **120** and **121** showed the CH_2 protons adjacent to the CN group, resonating as a singlet peak at 4.83 ppm. The IR spectra of these derivatives **120** and **121** showed strong absorption bands at 2165-2169 cm^{-1} assigned to the CN group.



Scheme 4.9 The synthesis of the acetonitrile intermediates **120** and **121**.

Synthesis of 119: The boronic ester **119** was synthesized starting from the commercially available 3-iodophenol **122** where it was dissolved in DMF with NaOH and the mixture was stirred until a pale brown solution was obtained (Scheme 4.10). Alkylation of the OH group was then achieved using the dropwise addition of 2-chloroacetonitrile **123** at room temperature and the mixture was stirred for 18 h. The mixture was then poured onto ice, where the resulting oil was extracted. Washing the oil extract with water and brine, and trituration with charcoal and concentration gave 2-(3-iodophenoxy)acetonitrile **124** in 80% yield. Analysis of the ^{1}H NMR spectrum of the intermediate **124** showed a singlet peak at 4.68 ppm assigned to the CH_2. Analysis of the EI-MS spectrum showed a peak at 259 *m/z* assigned to the molecular ion (M+).



Scheme 4.10 The synthesis of the acetonitrile boronic ester **119**.

The intermediate 2-(3-iodophenoxy)acetonitrile **124** was then converted to the boronic ester using Miyaura borylation[218] reaction, using bis(pinacolato)diboron **125**, potassium acetate and $PdCl_2(dppf)$* catalyst in THF and the mixture heated at reflux for 24 h. The solvent was then evaporated, ethyl acetate was added and the resulting residue was sonicated for 10 min and filtered. The filtrate was triturated with $MgSO_4$ and charcoal, filtered through celite and then concentrated under vacuum. The resulting residue was recrystallised from ethyl acetate/hexane (3:2) to give the boronic ester **119** as a pink solid in 95% yield. Analysis of the ^1H NMR spectrum of **119** showed a singlet peak resonating at 1.26 ppm integrating for 12 protons assigned to the four methyl groups. Analysis of the ^{13}C NMR spectrum showed one peak resonating at 83.7 ppm integrating for two carbons assigned to the dioxaborolan, and another peak at 25.2 ppm (four carbons) that was assigned to the methyl groups.

The Miyaura borylation reaction proceeds *via* the oxidative addition of the alkyl iodide to the palladium catalyst (Figure 4.6). The acetoxy anion from the base then reacts with the palladium complex replacing the iodide ion and forming the (acetate)palladium(II) complex. Transmetalation with the bis(pinacolato)diboron and subsequent reductive elimination then occurs to afford the boronic ester **119** and releasing the ligand complex.

* [1,1'-Bis(diphenylphosphino)ferrocene]dichloropalladium(II)

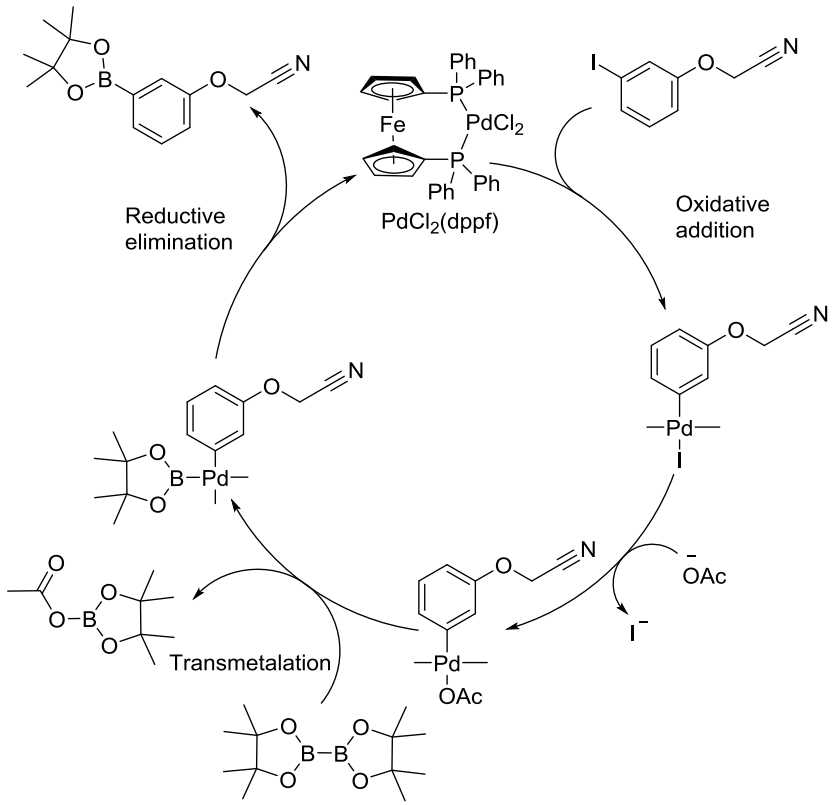
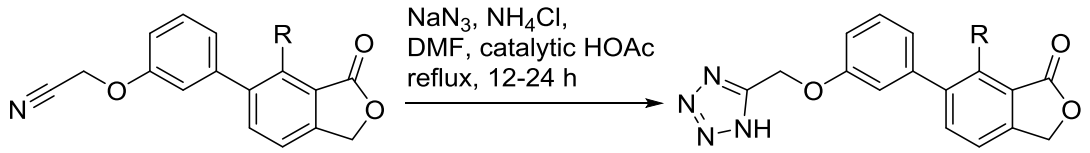


Figure 4.6 Borylation mechanistic pathway using $PdCl_2$(dppf) for the synthesis of the boronic ester **119**.[218]

The tetrazole ring synthesis: The benzolactone acetonitrile intermediates **120** and **121** were then used in the tetrazole synthesis step, utilizing the [2 + 3] cycloaddition of azides. The nitriles **120** and **121** were allowed to react with 4 equivalents each of NaN_3 and NH_4Cl in DMF and the mixture was heated at reflux for 48 h. DMF was then evaporated and the resulting residue was dried under vacuum. Ethanol was then added to the dry residue and the mixture was sonicated for 10 min, filtered and ethanol was concentrated and the residues were subjected to PLC using 90% CH_2Cl_2, 10% ethanol and 0.1% TFA. The isolated yields of the tetrazoles **126** and **127** were below 30%. Acid catalysis[219] for this cycloaddition significantly reduced the reaction time (Scheme 4.11). However, the improvement in the yield was not significant (35%). The purity of the isolated tetrazoles were analysed by HPLC and was found to be >95%.



120, R= OCH_3
121, R= CH_3

126, R= OCH_3, 34%
127, R= CH_3, 35%

Scheme 4.11 The synthesis of the tetrazoles **126** and **127** from the acetonitrile derivatives **116** and **117**.

Himo *et al*.[219] reported different mechanisms of tetrazole formation using azide salts in the presence of ammonium salts. The reaction can either proceed *via* the known [2 + 3] cycloaddition (pathway A, Figure 4.7) or by a two-step sequence where the azide, as a nucleophile, attacks the nitrile carbon giving an intermediate compound (pathway B, Figure 4.7), followed by ring closure through simple 1,5-cyclization. The second mechanistic pathway B depends of the substituent adjacent to the nitrile group (A in Figure 4.7), the more electron-withdrawing group, the intermediate in pathway B becomes more stable and mechanism B is more favourable based on calculated energy barriers[219] for the transition states. Also, the presence of a proton in the reaction medium can act as an activator to the nitrile group to be attacked. This mechanism is similar to the known mechanisms of acid-catalyzed nitrile hydrolysis and the Pinner synthesis,[220] with the difference being that the nucleophile in this case (Figure 4.7) is an azide.

The key to this mechanistic pathway (B) is that the nitrile is activated by a proton, provided by the ammonium salts as well as other proton sources. In this reaction (Scheme 4.11), the addition of acetic acid catalyst may have provided sufficient concentrations of the activating proton, and the reaction proceeded *via* pathway B, and that was the reason for shortening the reaction time (48 h without acid and 12-24 h in presence of acid).

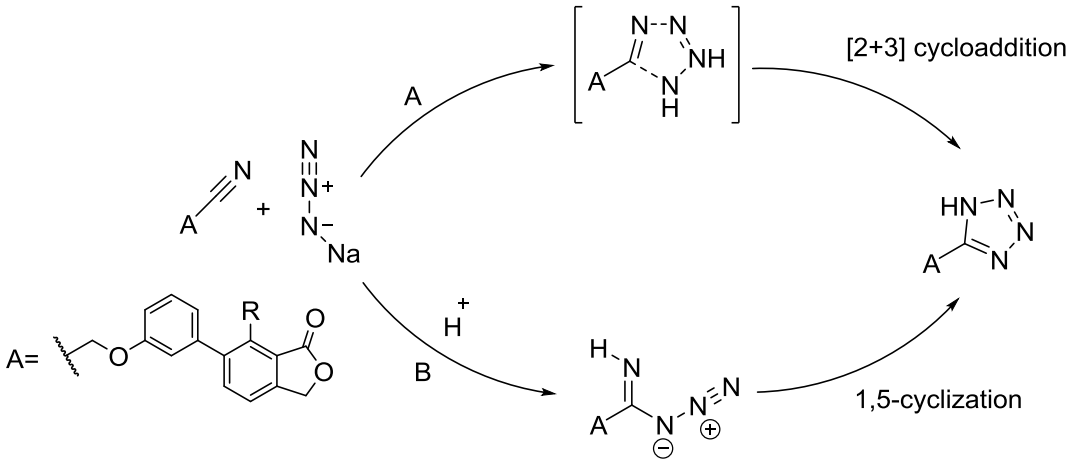
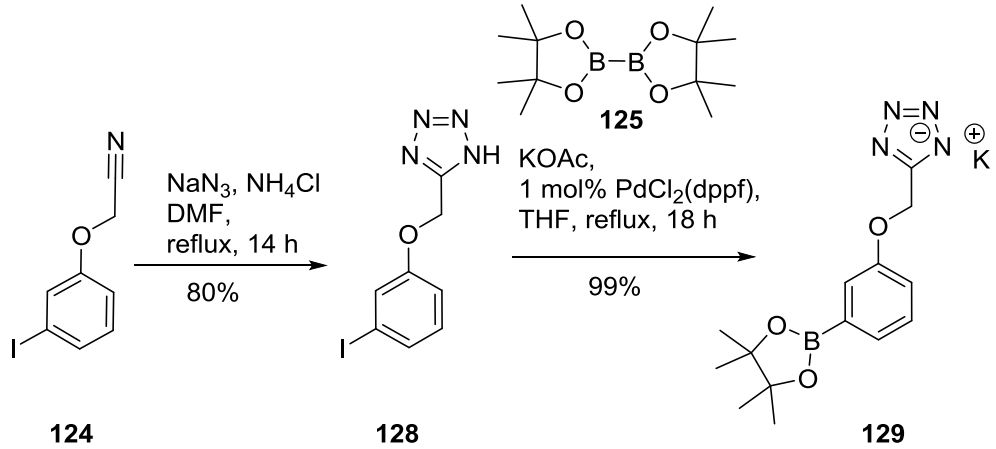


Figure 4.7 Mechanistic pathways for the tetrazole formation.[219]

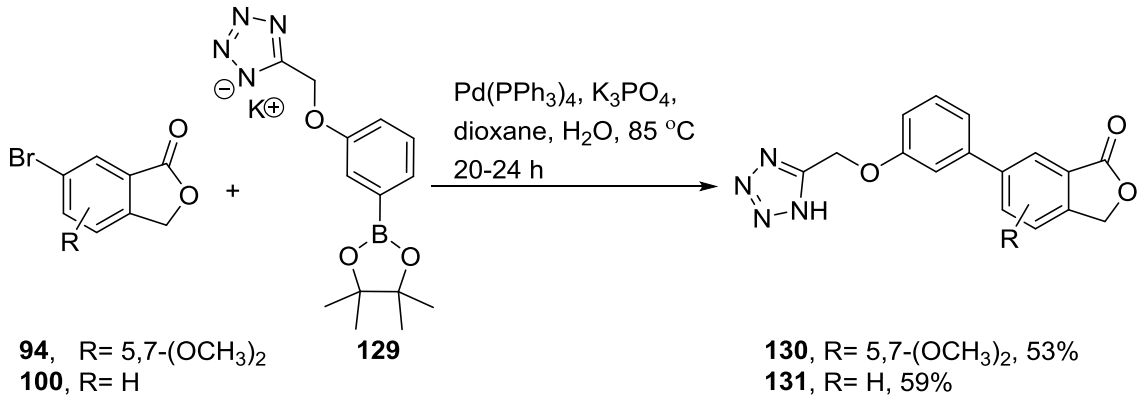
Access to the tetrazole via an alternative strategy: The modest yields (35%) of the tetrazole formation described above (Scheme 4.11) were not appropriate for the synthesis of other tetrazole derivatives. Another synthetic strategy was therefore designed to synthesize these derivatives in which the tetrazole moiety was attached to the boronic ester, prior to the coupling reaction with the 6-bromobenzolactones. The previously synthesized 2-(3-iodophenoxy)acetonitrile **124** was used to access the intermediate **128** utilizing the azide tetrazole formation discussed above. Therefore, **124** was mixed with NaN_3 and NH_4Cl in DMF containing a few drops of glacial acetic acid and the mixture was heated at reflux for 14 h under a N_2 atmosphere (Scheme 4.12). The solvent was then concentrated and the resulting residue was dried. Ethanol was added, and the mixture was sonicated for 10 min, filtered and ethanol was evaporated to yield a white fluffy solid of the tetrazole intermediate **128** in good yield (80%). The EI-MS spectrum for this tetrazole intermediate **128** showed a peak at 302 *m/z* assigned to the molecular ion (M+). This 5-((3-iodophenoxy)methyl)-1*H*-tetrazole **124** was then converted to the boronic ester using the same borylation method discussed in Scheme 4.10. Therefore, **128** was reacted with bis(pinacolato)diboron **125**, potassium acetate and $PdCl_2$(dppf) catalyst in THF and the mixture was heated at reflux temperature for 18 h under a N_2 atmosphere (Scheme 4.12). The solvent was then concentrated and ethanol was added to the resulting residue and the mixture was further heated at reflux for 10 min, filtered hot and concentrated under vacuum to afford a white solid of potassium 5-((3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)methyl)tetrazol-

1-ide **129** (Scheme 4.12). Analysis of the ^{1}H NMR spectrum of this boronic ester **129** showed a singlet peak resonating at 1.90 ppm and integrating 12 protons assigned to the four methyl groups. These four methyl carbons appeared also as a singlet peak in the ^{13}C NMR spectrum, resonating at 22.6 ppm. ESI-MS spectrum showed a peak at 301.1 *m/z* that was assigned to the anion $[m\text{-}K^{+}]^{-}$.



Scheme 4.12 The synthesis of the tetrazole boronic ester intermediate **129**.

The tetrazole synthesis: The synthesized boronic ester-tetrazole conjugate **129** was then reacted in the same optimized coupling conditions, using 1,4-dioxane/water mixture, with 6-bromo-5,7-dimethoxyisobenzofuran-1(3*H*)-one **94** and 6-bromoisobenzofuran-1(3*H*)-one **100** (Scheme 4.13). The tetrazole rings are known to be stable[198] in such mild coupling conditions as well as many other chemical conditions. Therefore, the mixture was heated at 85 ºC for 20-24 h, the solvent then was concentrated and a solution of $NaHCO_3$ was added to the resulting residue and sonicated for 15 min. The resulting turbid solution was then washed with CH_2Cl_2 and the aqueous layer was filtered, cooled to 0 ºC and neutralized with 2 M HCl. The resulting white solid was collected by filtration, washed with chilled methanol and dried under vacuum to afford the tetrazoles **130** and **131** in 53% and 57% yield, respectively (Scheme 4.13). The HPLC analysis of this series indicated purity >95%.

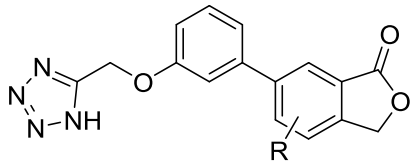


Scheme 4.13 Alternative route to access the tetrazole conjugates **130** and **131** *via* reaction with the boronic ester-tetrazole intermediate **129**.

Interestingly, using this alternative route (Scheme 4.13) not only allowed the use of mild reaction conditions, thereby avoiding unwanted side products, but also improved the yield of the final tetrazole conjugate (from 35% in Scheme 4.11 to 56% in Scheme 4.13) without further optimization. This reaction (Scheme 4.13) was clean compared to the above mentioned cycloaddition tetrazole formation method (Scheme 4.11) as judged by less side products spots upon TLC analysis.

For all the tetrazole conjugates obtained by the two methods, the sodium azide cycloaddition, **126** and **127** (Scheme 4.11) and the coupling with the tetrazole-boronic ester intermediate, **130** and **131** (Scheme 4.13), the purity was analysed using the normal phase HPLC, utilizing a mixture of isopropanol with 0.1% TFA in hexane. This system was developed after the failure with the reverse phase HPLC where the tetrazole conjugates were eluted first within two minutes, with poor resolution. Table 4.3 showed the calculated physiochemical properties for the synthesized tetrazoles.

Table 4.3 Physiochemical properties of the benzolactones-carboxylic acid conjugates.



R=5-CH_3, 5-OCH_3, 5,7$(OCH_3)_2$, H

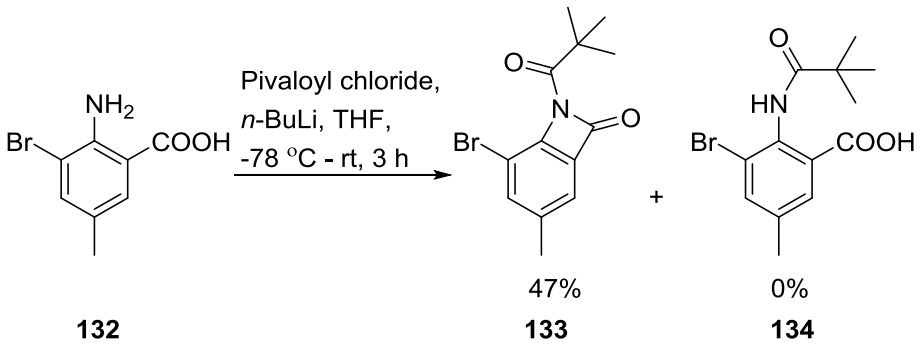
Compound	R	Mwt	clogP*	LogD† at pH		
				1	5.5	7.4
126	5-OCH_3	338.32	0.68±0.70	2.33	1.55	-0.16
127	5-CH_3	322.32	1.55±0.68	2.85	2.08	0.34
130	5,7-$(OCH_3)_2$	368.35	0.39±0.72	2.20	1.42	-0.29
131	H	308.30	1.09±0.68	2.47	1.69	-0.03

It can be inferred from table 4.3 that the designed and synthesized benzolactone-tetrazoles conjugates **126**, **127**, **130** and **131** have good ionization profiles when present at physiological pH. Like the synthesized carboxylic acid analogues **114-117**, the tetrazole conjugates **126**, **127**, **130** and **131** are nearly non-ionizable in the acidic pH of the stomach, however, their LogD values lie in the acceptable scale range (1-4).[186] The noticed increase in the lipophilicity compared to the carboxylic acid conjugates **114-117**, was related to the introduction of the more lipophilic tetrazole moiety. After absorption, the tetrazole conjugates **126**, **127**, **130** and **131** would tend to be strongly ionisable in the blood pH as can be indicated by the low LogD values. The methyl substituent **127** was the only derivative with a positive LogD value at pH 7.4. This argues for the increased blood solubility and also the possibility of conjugation with blood soluble proteins. However, at the intestine pH 5.5, the conjugates have relatively good LogD values (1.42-2.08), very similar to the carboxylic acid conjugates **114-117** (Table 4.2), which indicates the possible oral bioavailability of these compounds.

Synthesis of a β-lactam derivative: During the synthesis of the benzolactone derivatives, it was also of interest to synthesize one benzolactone derivative with an amino group (NH_2) replacing the OH group, which is responsible for the metabolic

* Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)
† Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA

drawback of the lead MPA **15** (Figure 4.1). Therefore, the synthesis of this amino analogue started with the 2-amino-3-bromo-5-methylbenzoic acid **132**. The benzolactone formation reaction (Scheme 4.1) required protection of the NH_2 group in the presence of palladium catalyst, and heating at elevated temperature for a long time would produce undesired side products that might involve the NH_2 group, and therefore, a base stable group such as the acetyl or the pivaloyl group, was required. Initial trials with mild conditions such as TMA in CH_2CH_2 were used, however, this method ended up with the full recovery of the unprotected amino benzoic acid **132**. Instead, an alternative method was used (Scheme 4.14) using pivaloyl chloride, excess *n*-BuLi as the base, in THF at -78 ºC. This method was previously described[221] for the protection of the unsubstituted 2-aminobenzoic acid. Therefore, the amino benzoic acid **132** was dissolved in THF and the solution was cooled to -78 ºC and was stirred for 20 min. Pivaloyl chloride was then added dropwise at -78 ºC and the reaction was allowed to warm to room temperature and was stirred for 3 h. The reaction mixture was then diluted with ethyl acetate and water and the organic layer was separated and washed with water and brine. The combined organic layers were concentrated under vacuum to yield a colourless oil (Scheme 4.14).



Scheme 4.14 Attempted reaction to synthesize the protected amino benzoic acid **134**.

Surprisingly, analysis of the resulting oil showed that the amino-protected compound **134** was not formed, instead, the β-lactam product **133** was obtained in 47% yield. The remaining of the reaction mixture was un-reacted starting material **132**. The EI-MS spectrum showed a peak at 295 *m/z* assigned for M+ of **133**, rather than a peak at 313 *m/z* for **134**. This 18 amu unit difference (as H_2O) in the molecular ions was an indication of the formation of the β-lactam ring system. The HRMS spectrum for the β-

lactam formula $C_{13}H_{15}BrNO_2$ (calculated; 296.0286 *m/z*), showed the ion peak at 296.0294 *m/z*, confirming the molecular formation of **133**.

This reaction proceeded *via* nucleophilic attack of the N atom on the carbonyl of the carboxylic acid group, forming the β-lactam ring. Figure 4.8 shows suggested mechanistic pathways for such reaction. Pathway (A) is the formation of the mixed acid anhydride, then the nucleophilic attack of the amino group on the carbonyl group, forming the closed ring, and then the less nucleophilic amidic NH attacked another pivaloyl group giving the final product **133**. The alternative pathway (B) is the protection of both functional groups, followed by the nucleophilic attack of the nitrogen on the electrophilic carbonyl and ring closure.

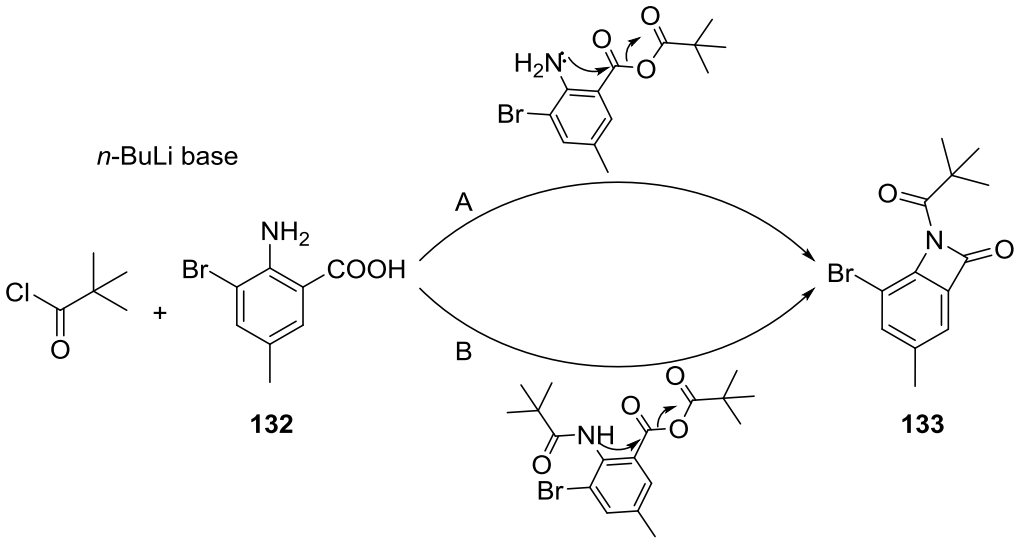


Figure 4.8 Suggested mechanistic pathways for the formation of the β-lactam product **133**.

A similar reaction was reported for the synthesis of such β-lactam (Figure 4.9) where the reaction occurred after protecting the NH_2 with the pivaloyl group.[222] Such evidence might suggest that the second pathway (B) might be the mechanism that occurred for **132**, especially the reaction (Scheme 4.14) occurred in the presence of excess base.

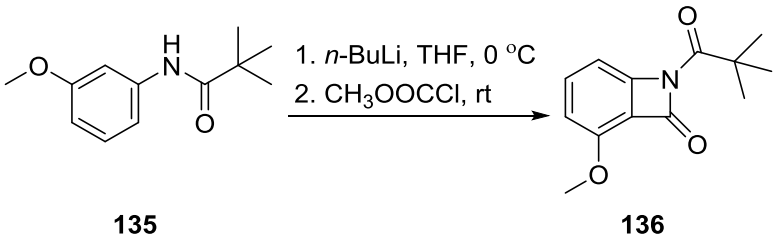
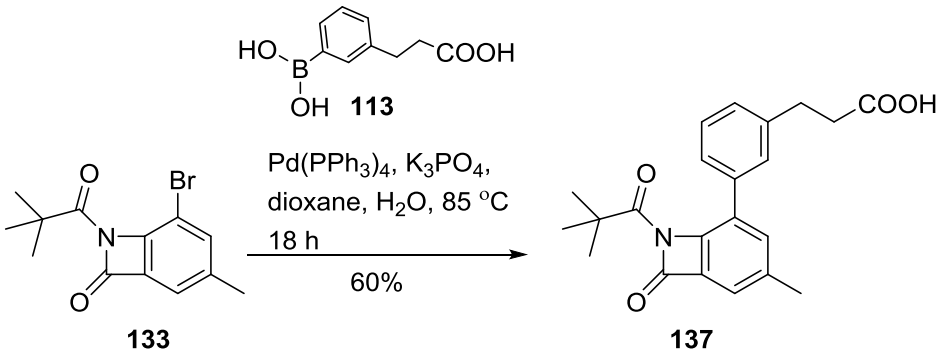


Figure 4.9 Reported reaction showing the formation of the β-lactam product **136**.[222]

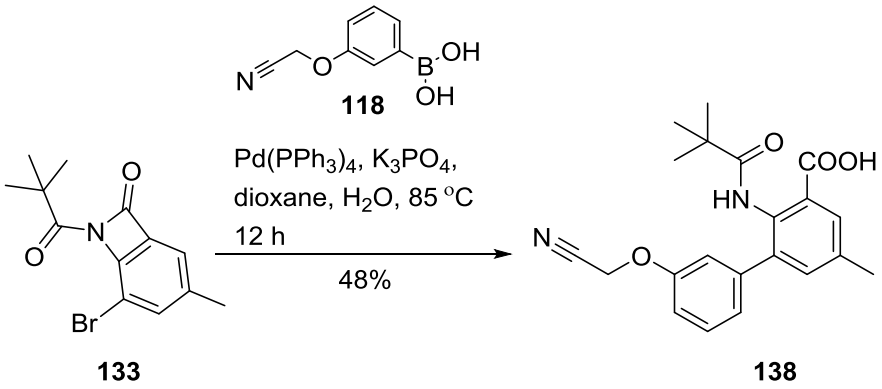
It was interesting to utilize this β-lactam derivative **133** in the coupling reactions with the different boronic acids, especially it has already the coupling handle (bromide) attached. Therefore, **133** was heated with 3-(3-boronophenyl)propanoic acid **113**, tetrakis(triphenylphosphine)palladium(0) and potassium phosphate in dioxane/water mixture, at 85 ºC for 18 h (Scheme 4.15). The solvent was then concentrated and the flask was cooled to 0 ºC, chilled HCl solution was added dropwise with stirring, the resulting solid was extracted with ethyl acetate, washed with water, brine and dried ($MgSO_4$), and the solvent was concentrated under vacuum. The resulting residue was subjected to a flash column chromatography using 50% CH_2Cl_2, 50% ethyl acetate and 1% acetic acid to afford 3-(3-(3-methyl-8-oxo-7-pivaloyl-7-azabicyclo[4.2.0]octa-1,3,5-trien-5-yl)phenyl)propanoic acid **137** as an oil in 60% yield. Analysis of the ^1H NMR of **137** showed a singlet peak at 8.56 ppm assigned to the COOH group. EI-MS spectrum showed a peak at 365 assigned to the molecular ion (M+). This oily product was found to be > 98% pure by HPLC analysis, with a retention time of 27.9 min.



Scheme 4.15 Coupling reaction of the β-lactam derivative **133** with the boronic acid **113**.

This β-lactam acid conjugate **137** has a LogD* value of 3.85 (at pH = 1), 2.83 (at pH = 5.5) and 0.99 (at pH = 7.4). The relatively unionized (clogP† = 3.75±0.65) form in the stomach indicated good lipid solubility that the compound can be found in the absorbable form. In the blood, **137** has the tendency to ionize providing good solubility and therefore, good distribution within the body. The slight ionization state in the jejunum (pH=5.5) makes **137** also a possible orally bioavailable candidiate.

The reaction of the β-lactam intermediate **133** with the boronic acid (3-(cyanomethoxy)phenyl)boronic acid **118** was undertaken utilizing the same coupling conditions (Scheme 4.16). The coupling reaction occurred, forming the biphenyl system, and the β-lactam was open to give a benzoic acid derivative **138**. The IR spectrum of **138** showed a broad peak at 3328 cm^{-1} assigned to the NH group. Analysis of the ^1H NMR spectrum showed a peak resonating at 8.61 ppm that was assigned to the COOH group. The found mass by the HRMS was 367.1656 that corresponded to the formula $C_{21}H_{23}N_2O_4$ (MH+; calculated, 367.1658).



Scheme 4.16 Reaction of the β-lactam **133** with (3-(cyanomethoxy)phenyl)boronic acid **118**.

The difference between this reaction (Scheme 4.16) and the previous one (Scheme 4.15) was only during the workup, which was the same except the HCl was not pre-cooled and was added at room temperature. It is believed that during this thermally uncontrolled procedure, the β-lactam ring was hydrolysed to give the COOH group of **138**. The LogD profile for this compound is shown in Table 4.4.

* Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA
† Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)

Table 4.4 Calculated logP and LogD for the open β-lactam product **138**.

clogP* (unionized)	LogD†		
	pH 1	pH 5.5	pH 7.4
4.14±0.74	4.6	3.13	1.30

The presence of a carboxylic acid group within the structure is responsible for the unionized state of the compound in acidic pH environment (stomach), and therefore, would be present as a lipophilic state that can pass the gut wall. After reaching the blood (pH 7.4), the compound ionizes providing blood solubility. At pH 5.5 (jejunum), where absorption occurs, the LogD value is within the acceptable range (1-4)[186] for orally bioavailable drugs.

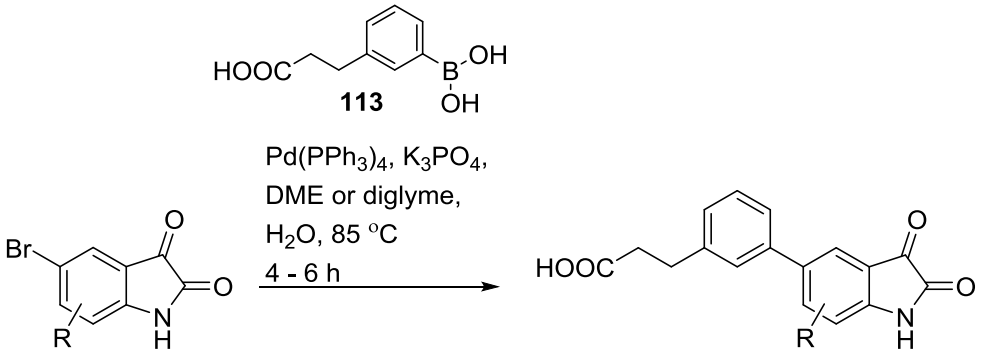
4.4.2. Coupling with the isatin aromatic head groups

4.4.2.1. Coupling to the linker and COOH tail

Synthesis of the 5-aryl isatin derivatives starting from the corresponding 5-halo-isatin *via* the cross-coupling Suzuki reaction has been previously described,[223,224] using tetrakis(triphenylphosphine)palladium(0) as a catalyst and potassium phosphate as a base. Therefore, the 5-bromoisatin derivatives (**103**, **106**, **112** and the 5-bromoindoline-2,3-dione **139**) were reacted with the 3-(3-boronophenyl)propanoic acid **113**, tetrakis(triphenylphosphine)palladium(0) and potassium phosphate (Scheme 4.17) in either DME or diethylene glycol dimethyl ether (diglyme) based on the solubility of the 5-bromoisatin used: in case of 5-bromo-4,6-dimethoxyindoline-2,3-dione **103** and 5-bromo-4-methoxy-7-methylindoline-2,3-dione **112**, the solvent used was diglime. The reaction was then heated at 85 ºC under nitrogen atmosphere for 4-6 h. After cooling to room temperature, the solvent was concentrated and HCl solution was added, and the resulting suspended solid was either collected by filtration or extracted with ethyl acetate. The combined residues were then dissolved in ethanol and were subjected to PLC (10% methanol in CH_2Cl_2, 0.5% acetic acid) to yield the isatin-acid conjugates **140**-**143**. Analysis of the ^{1}H NMR spectrum for 3-(3-(2,3-dioxoindolin-5-

* Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)
† Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA

yl)phenyl)propanoic acid **143** showed two triplets resonating at 2.88 and 2.59 ppm, both integrating for two protons that were assigned to the two CH_2 groups introduced from the boronic acid. Analysis of the IR spectrum of **143** showed an absorption band at 2989 cm^{-1} assigned to the OH of the COOH group. The purity of this series was >95% as analysed by HPLC, with retention times 5.34-35.84 min.



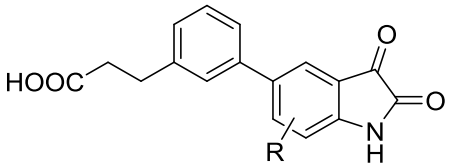
103, R= 4,6-$(OCH_3)_2$
106, R= 7-CH_3
112, R= 4-OCH_3-7-CH_3
139, R= H

140, R= 4,6-$(OCH_3)_2$, 40%
141, R= 7-CH_3, 55%
142, R= 4-OCH_3-7-CH_3, 48%
143, R= H, 65%

Scheme 4.17 The synthesis of the isatin-acid conjugates **140-143**.

Table 4.5 shows the calculated physical properties of this series of compounds **140-143**. For all the derivatives, the ionization increases as the pH increases, the LogD values at pH 5.5 ranges from 1.13 to 1.71, a good range for oral bioavailability. However, the increased solubility in the blood pH, as indicated by the negative LogD values, might argue the chance of binding with the blood soluble proteins.

Table 4.5 Physiochemical properties of the isatin-carboxylic acid conjugates



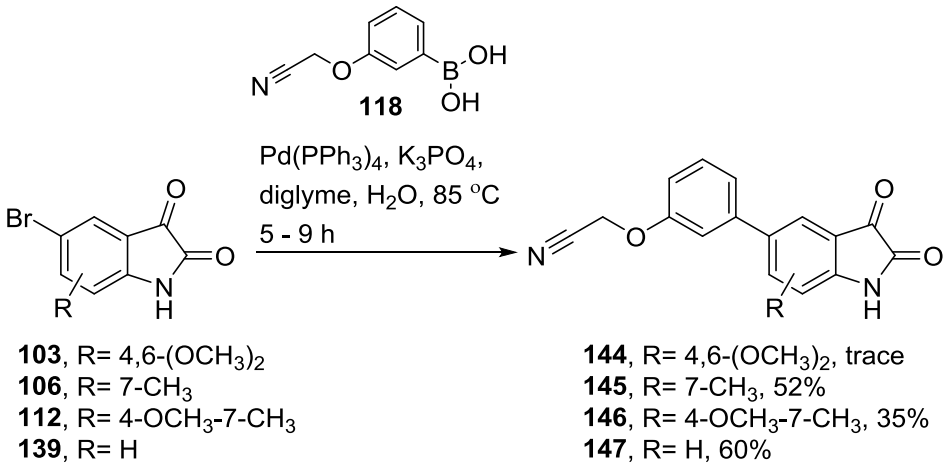
R= H, 7-CH_3, 4-OCH_3-7-CH_3, 4,6-$(CH_3)_2$

Compound	R	Mwt	clogP*	LogD† at pH		
				1	5.5	7.4
140	4,6-$(CH_3)_2$	355.35	1.49±0.82	2.14	1.13	-0.70
141	7-CH_3	309.32	2.40±0.37	2.53	1.55	-0.28
142	4-OCH_3-7-CH_3	339.35	2.06±0.80	2.70	1.71	-0.12
143	H	295.29	1.94±0.37	2.17	1.19	-0.64

3.4.2.2. Coupling to the linker and tetrazole tail

In order to obtain the isatin-tetrazole conjugates, the acetonitrile derivatives were synthesized by reacting the 5-bromoisatin derivatives (**103**, **106**, **112** and **139**) with (3-(cyanomethoxy)phenyl)boronic acid **118** using diglyme/water mixture as a solvent in the presence of tetrakis(triphenylphosphine)palladium(0) and potassium phosphate and the mixtures were heated at 85 ºC for 5-9 h (Scheme 4.18). The solvent was then concentrated and 1 N HCl solution was added, and the suspended solids were either collected by filtration or extracted with ethyl acetate. The combined residues were adsorbed onto silica and subjected to flash column chromatography (40% ethyl acetate in petroleum spirit) to yield the acetonitrile intermediate **144-147** in 35-60% yield. The IR spectrum of 2-(3-(2,3-dioxoindolin-5-yl)phenoxy)acetonitrile **147** showed an absorption band at 2166 cm^{-1} that was assigned to the CN group. Analysis of the ^1H NMR spectrum of **147** showed a peak resonating at 5.21 ppm and integrating for two protons that was assigned to the CH_2 group.

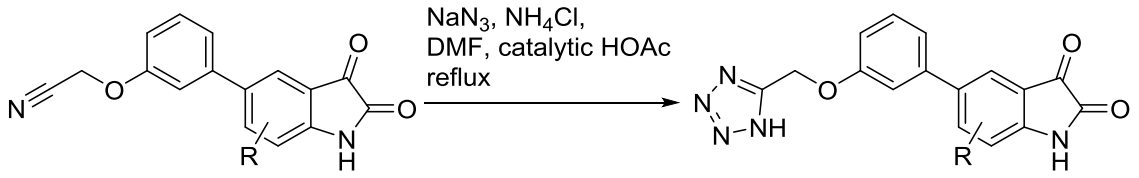
* Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada)
† Calculated using PALLAS3.7.1.2 CompuDrug Chemistry Ltd, USA



Scheme 4.18 The synthesis of the isatin-acetonitrile intermediates **140**-**143**.

Among the isatin-acetonitrile intermediates **144**-**147**, 2-(3-(4,6-dimethoxy-2,3-dioxoindolin-5-yl)phenoxy)acetonitrile **144** was not isolated from the crude mixture due to product instability. The EI-MS spectrum of a sample taken from the reaction mixture showed a peak at 338 that assigned to the molecular ion (M+). The HRMS spectrum showed a peak at 339.0980 *m/z* that was assigned to the formula $C_{18}H_{15}N_2O_5$ for (MH+; calculated, 339.0981 *m/z*). The reaction for **144** was repeated several times in order to isolate a pure compound, however, once the solid was separated from the reaction mixture, it was found to decompose to many overlapping spots on the TLC plate. Subjecting the solid to a flash column chromatography was not successful and analysis of the EI-MS spectra of the isolated fractions did not show any indications of **144**.

Conversion to tetrazoles: The reaction of the isatin-acetonitrile derivatives **145-147** with sodium azide and ammonium chloride was attempted to afford the isatin-tetrazole conjugates (Scheme 4.19). Several attempts were made using different molar ratios of the azide/amonium chloride up to 10 equivalents, and heating at elevated temperatures for longer times up to 72 h with and without acid catalysis, however, the tetrazole conjugates **148**-**150** were not formed. Upon workup and separation of the oily gummy material from the reaction, TLC analysis showed that the starting material disappeared and just a polar polymeric spot was found. No evidence indicating the formation of the tetrazole conjugate was observed in the mass spectra of these polymerized residues.



145, R= 7-CH_3
146, R= 4-OCH_3-7-CH_3
147, R= H

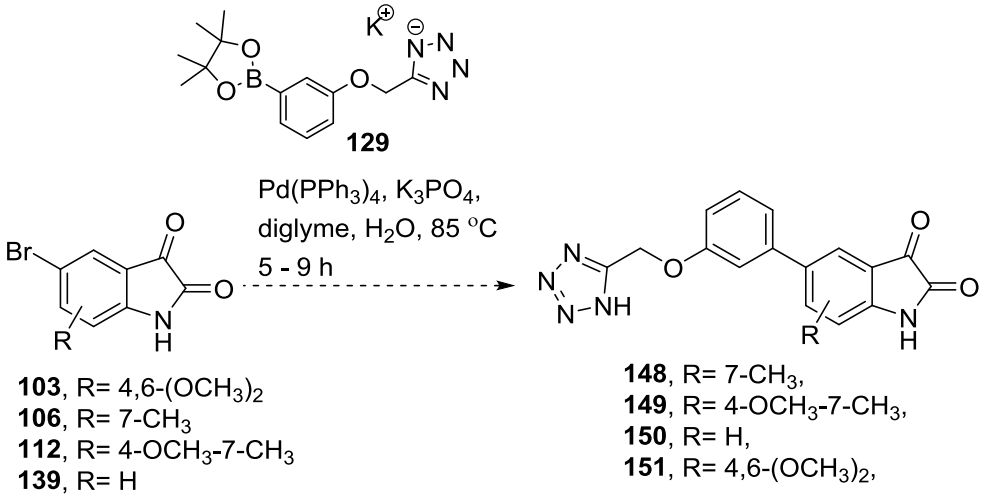
148, R= 7-CH_3, 0%
149, R= 4-OCH_3-7-CH_3, 0%
150, R= H, 0%

Scheme 4.19 Attempted access to the isatin-tetrazole conjugates **148-150** *via* the cycloaddition reaction with sodium azide/ammonium salts.

Such reaction of the isatin derivatives with the sodium azide has not been previously reported. The nucleophilicity of the sodium azide and the electrophilic C3 of isatin might be involved in such reaction, however, no evidence was found in the mass spectra to support this hypothesis. This reaction was attempted before the design of the alternative coupling reaction (Scheme 4.13) with the synthesized tetrazole-boronic ester salt **129**, and therefore, it would be interesting to try to access the isatin-tetrazole conjugates **148-150** utilizing such coupling reaction, in which the electrophilicity of the isatin C3 will not interfere with the reaction as in the case of sodium azide.

4.5. Concluding remarks

This chapter discussed the synthesis of a series of benzolactone-carboxylic acid conjugates **114-117**, benzolactone-tetrazole conjugates **126**, **127**, **130**, **131** and isatin-carboxylic acid conjugates **140-143** as analogues of the lead compounds mycophenolic acid **15**. The designed and synthesized analogues carry the same features as the parent lead compound in that they all have the three components: the aromatic head, the rigid linker and the acidic tail. All of the synthesized compounds showed good physiochemical properties such as the low molecular weight and good LogD profiles, having a range of 1-4 in the pH 5.5. Therefore, this series can be further developed as promising leads against CHIKV depending on the activities in the preliminary anti-chikungunya testing. The formation of isatin-tetrazole conjugates **148-151** might be possible using the coupling with the designed and synthesized boronic ester **129** (Scheme 4.20).



Scheme 4.20 Possible coupling reaction to access the isatin-tetrazole conjugates **148-151** *via* reaction with the boronic ester **129**.

A facile and applicable synthetic strategy was developed to access different analogues for the mycophenolic acid **15** *via* the introduction of the widely used cross coupling reactions. The obtained carboxylic acid conjugates can be further derivatised *via* the carboxylic acid reactions such as formation of esters or amides as prodrugs. The tetrazole conjugates can also be further derivatised *via* the substitution on the tetrazole NH by different alkylation/arylation reactions. The isatin conjugates can also be further derivatised utilizing the different nucleophilic reactions on C3 or ring expansion reactions (Figure 4.10). This will allow the further investigate the structure activity relationship of this series against the CHIKV.

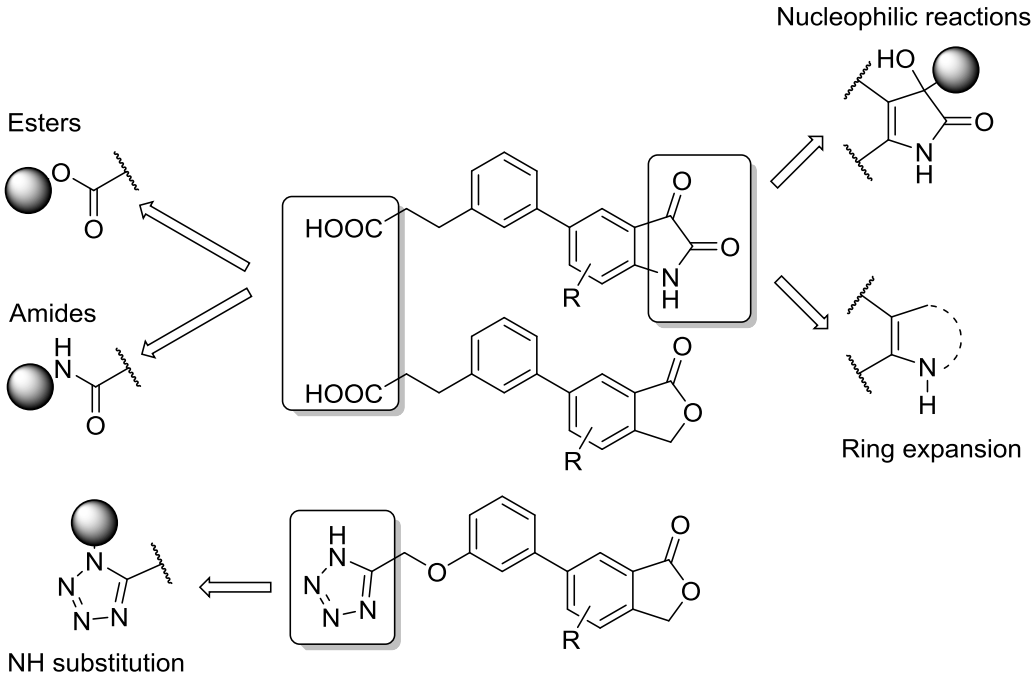


Figure 4.10 Possible derivatisation of the synthesized conjugates.

CHAPTER 5: Computational and Modelling Studies of the CHIKV nsP2 as a Possible Drug Development Target

5.1. Introduction

The CHIKV non-structural protein 2 (nsP2) is a significant enzyme that plays an important role during the viral replication cycle. It consists of 324 residues with two domains; the N and C domains that have different cellular functions. The nsP2 of *alphaviruses* is a multifunctional protein.[57,58,59,60] The N-terminus (amino terminal) was found to be the RNA triphosphatase that performs the first of the viral RNA capping reactions and a nucleotide triphosphatase (NTPase) that fuels the RNA helicase activity performed by the other domain.[65] It also performs 5`-triphosphatase (RTPase) activity that removes the γ-phosphate from the 5` end of RNA. Both NTPase and RTPase activities are completely dependent on Mg^{2+} ions.[65] The proteolytic domain has been allocated to its C-terminal (carboxy terminal) section which forms a papain like cysteine protease (also known as thiol protease).[57,61] The nsP2 proteolytic activity is critical for virus replication and responsible for the cleavages of the non-structural polyprotein complex during the viral life cycle.[62,63]

During the life cycle, and after entering the host cell, two non-structural protein precursors are translated from the viral mRNA, and are then cleaved generating nsP1, 2, 3 and 4. During translation, nsP123 binds to free nsP4, with some cell proteins forming the replication complex,[38,39] which synthesizes a full-length negative-strand RNA intermediate required for replication. When the nsP123 concentration increases, it is cleaved into nsP1, nsP2, nsP3 and nsP4 which form, along with host cell proteins, the positive strand replicase, which produces the 26S sub-genomic positive strand RNAs and genomic (49S) RNAs.[39] The proteolytic activity of the CHIKV nsP2 is responsible for the cleavage of the replication machinery into individual working proteins. The *alphaviruses* nsP2 proteins have also been described as virulence factors responsible for the transcriptional and translational shutoff in infected host cells and the inhibition of interferon (IFN)-mediated antiviral responses contributing to the controlling of translational machinery by viral factors.[72,73,225] This controlling comes through interactions with cellular RNA binding proteins, including heterogeneous nuclear

ribonucleoproteins (hnRNPs), ribosomal protein S6 (RpS6), and cellular filament components. Recently discovered, 22 cellular components are believed to interact with nsP2 or nsP4, contributing to the CHIKV replication, such as heterogeneous nuclear ribonucleoprotein K (hnRNP-K) and ubiquilin 4 (UBQLN4). It was also noted that the interaction of nsP2 with the tetratricopeptide repeat protein 7B (TTC7B) plays a significant role in the cellular machinery control induced by the CHIKV infection.[74] These critical functions are associated with nsP2 C domain protease and also the N domain that controls some metabolic processes of some cellular substrates, and therefore, this enzyme represents an important target in developing drugs against the CHIKV virus.

Both N and C domains are composed of α-helices and β-strands (Figure 5.1). The N terminus is dominated by α-helices (red spirals in Figure 5.1) whereas, the C-terminal domain contains helices and strands (cyan colour in Figure 5.1). The central β-sheets are flanked by α-helices. The crystal structure of CHIKV nsP2 protease was solved in 2012. Being a cysteine protease enzyme (within domain C), the catalytic mechanism involves a nucleophilic cysteine thiol in a catalytic dyad.[66] Analysis of the CHIKV nsP2 crystal structure shows 6 cysteine residues; three in the N-terminus (Cys1013, Cys1057 and Cys1121) and three in the C-terminus (Cys1233, Cys1274 and Cys1290) as shown in Figure 5.1. Since the proteolytic activity is isolated in the C-terminus,[64] one of the three cysteine residues in the C-domain might be the catalytic head.

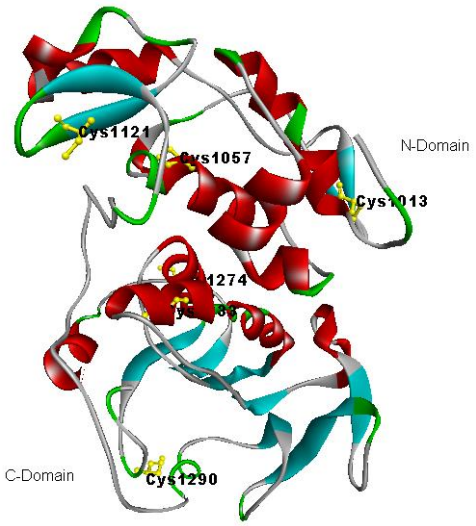


Figure 5.1 CHIKV nsP2 crystal structure showing the N- and C-terminal domains, cysteine residues shown in yellow balls and stick form (PDB code: 3TRK[67], no reference cited thus far).

Generally, the first step in the mechanism of cysteine protease catalysis is the deprotonation of a thiol group within the enzyme active site by an adjacent amino acid containing a basic side chain, usually a histidine residue.[68] Among the three cysteine residues in the C-terminus (Figure 5.2), the Cys1274 residue is less likely to be involved in the catalytic mechanism as only one His residue (His1314) is nearby, whereas for the other cysteine residues, four histidine residues (His1222, His1228, His1229 and His1236) are nearby and could be associated in the deprotonation mechanism (Figure 5.2).

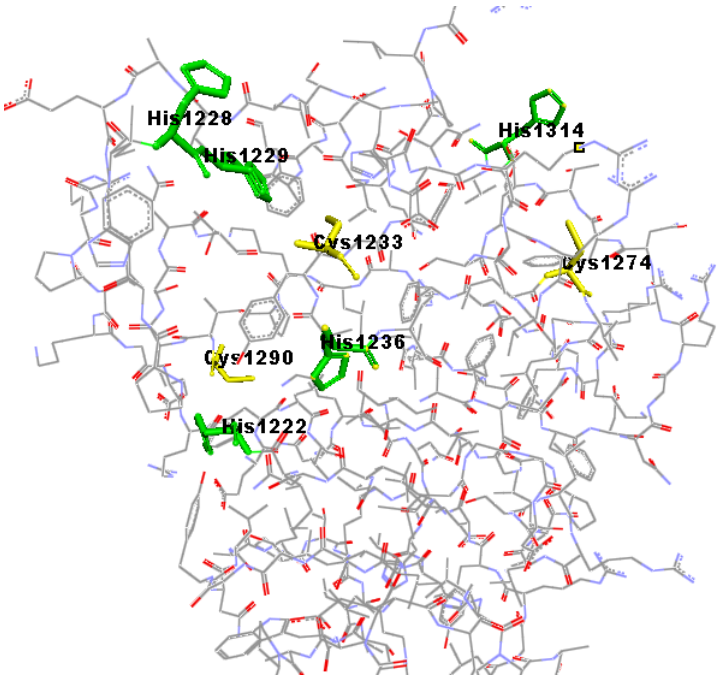


Figure 5.2 Schematic representation of the CHIKV nsP2 C-Domain showing the positions of the cysteine residues (in yellow) and histidine residues (in green), generated from the crystal structure PDB file code: 3TRK.

For the nsP2 protease activity, the enzyme must hold its substrate at the catalytic site to start the cysteine cleavage mechanism. This site can be described as the enzyme "mouth", and will be critical for the enzyme's function. Blocking this site, will therefore, antagonise the enzyme function and will interfere with the viral replication cycle inside the infected host cell.

In 2012, Singh Kh *et al*. reported the development of a homology model of CHIKV nsP2 protein based on the crystal structure of the nsP2 protein of Venezuelan equine encephalitis virus (VEEV),[69] before the release of the actual nsP2 crystal

structure. The critical residues in nsP2 were identified by docking three different peptides in order to identify the residues responsible for non-structural protein cleavage; nsP1-2, nsP2-3 and nsP3-4 peptides. These three peptide sequences represent the substrates for the nsP2 proteolytic processing with a remarkable preference of nsP3-4>nsP1-2>nsP2-3.[70] The active site was investigated and was found to lie in the C-terminal domain.[69] The key residues Gln1039, Lys1045, Glu1157, Gly1176, His1222, Lys1239, Ser1293, Glu1296 and Met1297 were found to interact with the nonstructural protein sequence complex to be cleaved, and were considered an individual functional unit. Only two residues are located in the N-domain; Gln1039 and Lys1045; all the other residues are located in the C-domain. In similar work, Bassetto *et al.*[71] reported the identification of *in silico* CHIKV nsP2 inhibitors based also on a homology model of CHIKV nsP2 protein, not the actual crystal structure.

Blocking the protease activity will be an applicable strategy for designing inhibitors for that particular enzyme. However, the reported studies[69,71] only included the C domain as the main target site for the enzyme. They also did not describe the active proteolytic site in sufficient detail, nor did they discuss the possibility of targeting sites in the N domain. Herein, we describe the protease active site revealing some important criteria for the future design of inhibitors. We also describe here for the first time the binding site within the nsP2 N domain. Blocking the later site might not only block the multiple functions of the N domain,[65] but could affect the C protease domain through an allosteric effect, invoking indirect inhibition of the C domain.

5.2. Results and discussion

The possible druggable sites at both the C and N domains of the CHIKV nsP2 were investigated. These sites were then used for the identification of possible inhibitors through virtual screenings.

5.2.1. *Protease active site (domain C)*

From the CHIKV nsP2 crystal structure analysis, one major cavity can be detected within the enzyme surface lying as a central region between domains C and N. It also blocks the access through the two of the cysteine residues within domain C; Cys1233, Cys1290. Therefore, this cavity is more likely to be the protease active site (Figure 5.3), where the protein holds the peptide segment to be cleaved.

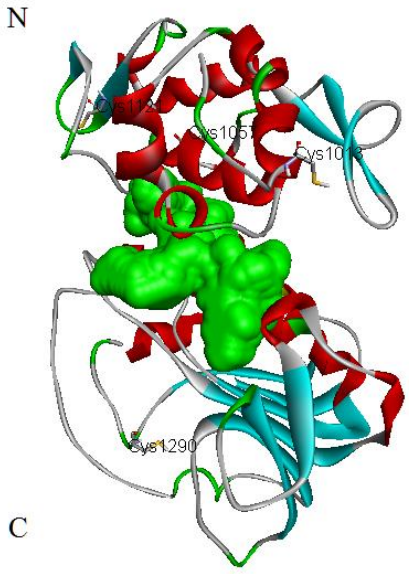


Figure 5.3 Crystal structure of the CHIKV nsP2 showing the enzyme mouth cavity in green space filling lying in the centre between domain N and C.

A closer inspection of the binding site revealed that the cavity looks as a U-shaped cavity or tunnel (hereafter called site_1) that extends inward from the enzyme surface and opens and widens towards the outer surface of the enzyme (Figure 5.4). The tunnel itself is narrow and is surrounded by side walls and floor on the exterior side of the enzyme.

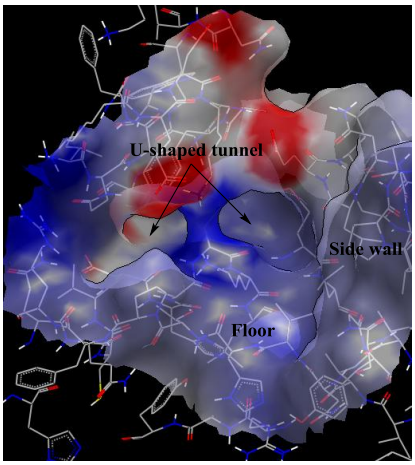


Figure 5.4 Site_1 U-shaped cavity that is lying between the nsP2 domains N and C. It has the two openings on the enzyme surfaces. The cavity mouth is wide and surrounded by hydrophobic side wall and a floor.

Site_1 makes contact with the residues: Gln1039, Ala1040, Lys1045-Tyr1047, Gly1176, Pro1191, Leu1192, Tyr1201, Asn1202, Leu1203, Glu1204, His1222, Thr1223, Pro1224, Arg1226, Asp1235, Mes1238, Lys1239, Leu1243. The cavity is mainly hydrophobic due to the presence of the residues Ala, Tyr, Gly, Pro and Lys, with some amino acids able to form H-bonds such as His, Tyr, Gln, Glu as well as the polar part of Lys. The central blue bridge (Figure 5.4) in the middle of the cavity is a hydrophilic bridge that is formed by the polar heads of the N domain Lys1045 and the C domain Lys1239. Domain C residue Cys1233 is buried in one of the α-helices behind this cavity and closer to domain N, confirming the hypothesis that this mouth cavity might be the access to the catalytic proteolytic activity of the enzyme. Therefore, blocking this cavity might block the proteolytic activity of the enzyme and consequently inhibit viral replication.

5.2.2. *N domain binding site*

Lying on the opposite side of the enzyme surface is another hydrophobic pocket that extends within the N domain of the enzyme before the junction with domain C (Figure 5.5).

a b

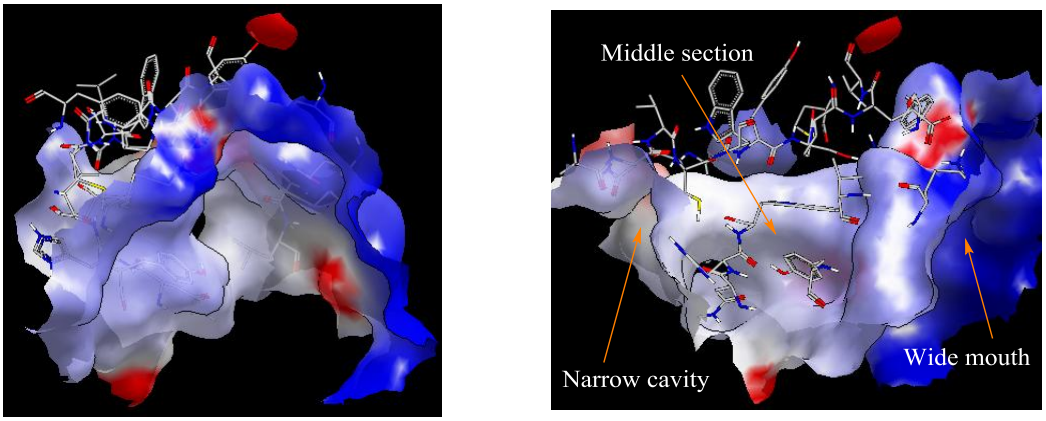


Figure 5.5 Hydrophobic pocket within the N domain of the CHIKV nsP2 a) Front view of the pocket, b) side view of the pocket, divided into 3 sections; the narrow linker cavity which open towards site_1 side, a middle section and a wide terminal mouth.

This pocket (hereafter called site_2) extends under the N domain β-sheet and ends with a small cavity that opens on the domain C side surface, just above site_1 (Figure 5.6).

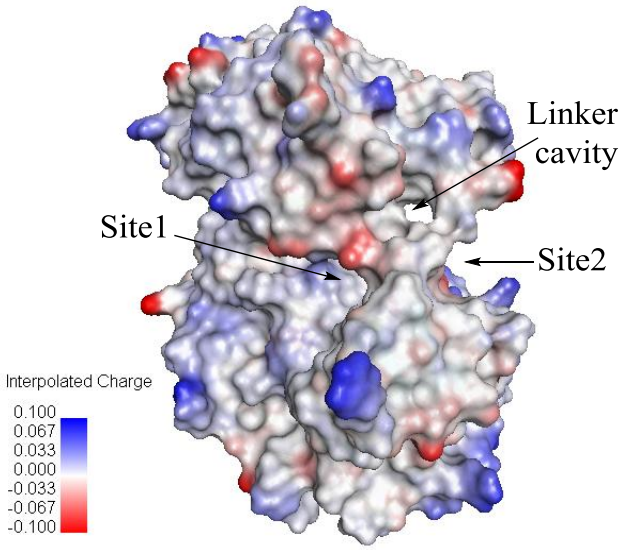


Figure 5.6 Electrostatic surface of the CHIKV nsP2 showing site_1, site_2 and the linker cavity that opens from site_2 towards site_1

Site_2 is mainly lined by residues: Asn1011, Cys1013, Trp1014, Ala1046, Tyr1047, Ser1048, Glu1050, Val1077, Tyr1079, Asn1082, Trp1084, Gly1090, Lys1091, Leu1205 and Mse1242. It is also a hydrophobic cavity due to the presence of hydrophobic residues with some amino acids able to form H-bonds such as Ser, Tyr, Trp, Gln, and also the polar moiety of Lys. Targeting this site by inhibitors, might not only block the NTPase and RTPase N domain functions, but may also act as an allosteric site for the protease binding site on the opposite C domain side of the enzyme; small molecules bound to this site may alter the function of the enzyme or change the enzyme conformation and consequently block the enzyme function, stopping the viral replication.

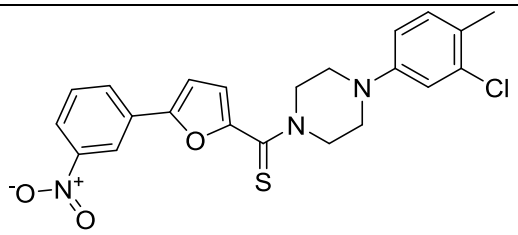
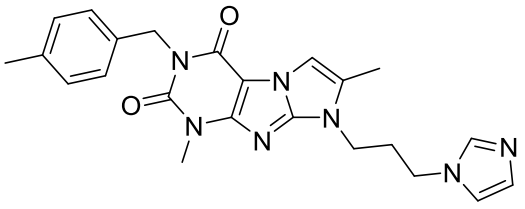
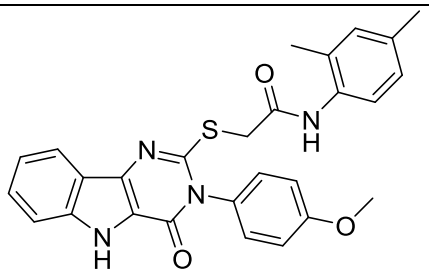
5.3. Virtual screening with the CHIKV nsP2

A virtual screening study was performed on both sites (site_1, site_2) using the Life chemicals cysteine protease inhibitors library (28,960 compounds). This library was designed using a ligand based approach – first, a set of 585 compounds active in assays

related to cysteine proteases was assembled. This Life Chemicals collection was searched for compounds similar to the reference dataset using MDL public keys and the Tanimoto similarity cut-off of 85% to generate a library of 28,960 compounds.

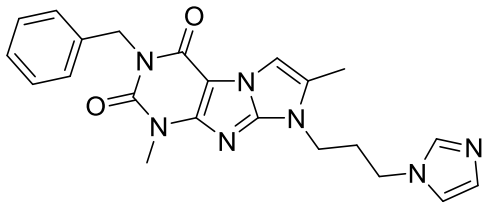
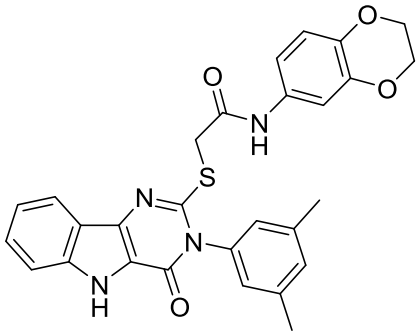
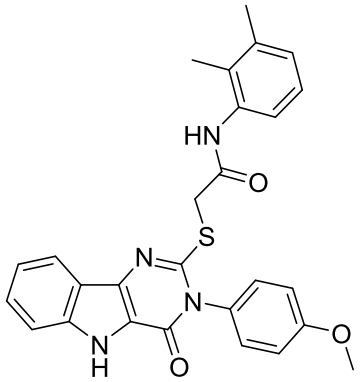
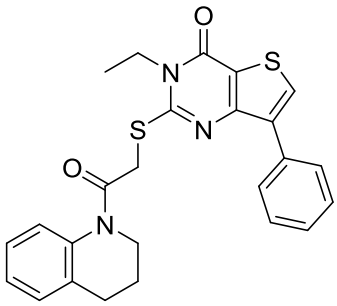
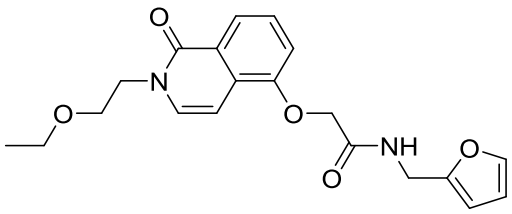
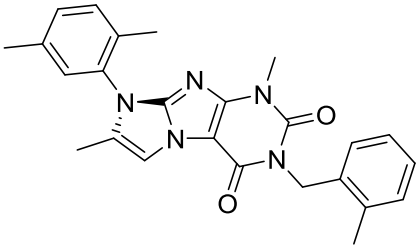
The virtual screening was performed using two successive docking scores, FRED (fast rigid exhaustive docking) for fast precise screening, with the top hits then subjected to a ranking scoring using the AUTODOCK algorithm. The important residues within both sites have been identified. Some *in silico* inhibitors were also predicted that could be developed as selective CHIKV nsP2 inhibitors. The results of the first Fred screening and the second Audodock re-ranking on both sites are shown in tables 5.1-5.6.

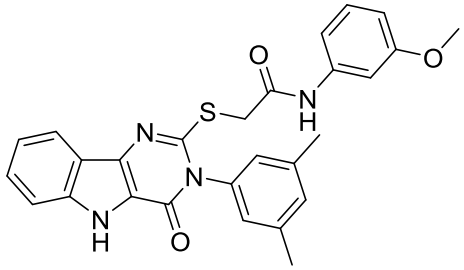
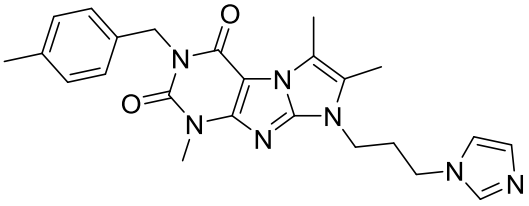
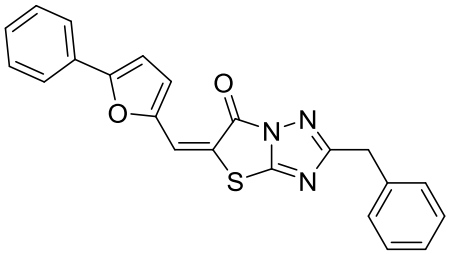
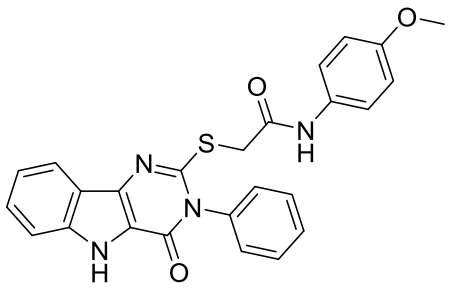
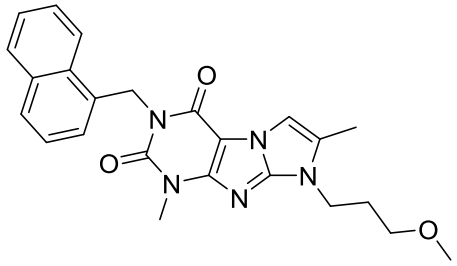
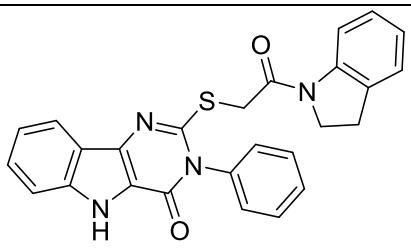
Table 5.1 The top 25 compounds docked by Fred in **nsP2_site_1** and their interacting residues. Numbering is based on Fred output list (visualised by VIDA[*]).

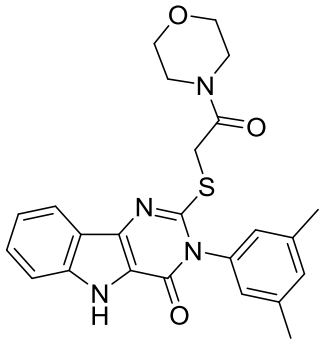
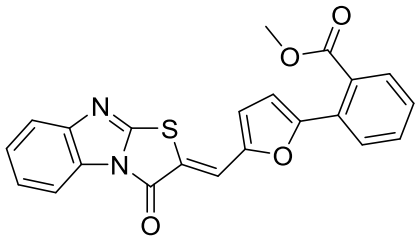
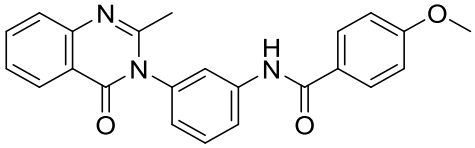
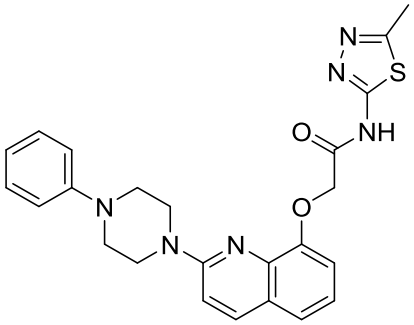
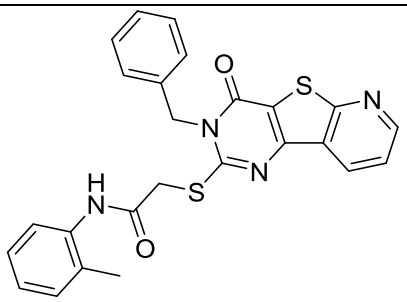
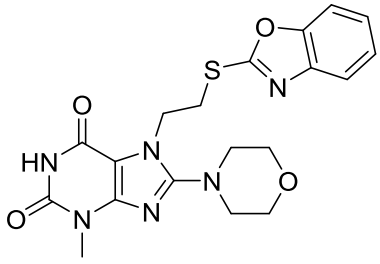
Entry	Structure	H-bonding residues	Other interacting residues
1[†]	Mg^{2+}	-	-
2		Lys 1239 Asp 1235	π-cation Lys 1045
3			3 x π-cation Lys 1239
4		Lys 1045	Lys 1239

[*] VIDA is a visualising software developed by Openeye (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com)

[†] Mg ion within the crystal structure

5			π-cation Lys 1239 π-sigma Lys 1239
6			2x π-cation Lys1239
7		Leu 1203	π-cation Lys 1239 π-sigma Lys 1239
8			π-cation Lys 1045 π-cation Lys 1239 π-sigma Lys 1239
9		Lys 1045 x2 Asp 1235	π-cation Lys 1239 x 2 π-sigma Lys 1239
10			π-cation Lys 1239 x 3

11			π-cation Lys 1239 x 2 π-sigma Lys 1239 π-cation Lys 1045
12			π-cation Lys 1239 x 2 π-sigma Lys 1239
13		Lys 1045	π-sigma Ile 1221 π-cation Lys 1045 π-cation Lys 1239 π-sigma Lys 1239 π-sigma Ala 1040
14		Lys 1045	π-cation Lys 1239 x 2
15			π-cation Lys 1239 π-sigma Lys 1239
16			π-cation Lys 1239 π-sigma Lys 1239

17			π-cation Lys 1239 x 2 π-sigma Lys 1239
18			π-cation Lys 1239 x 2 π-sigma Lys 1239
19			π-π His 1222 π-cation Lys 1239 π-sigma Lys 1239
20		Lys 1045 x2	π-cation Lys 1239 x 3
21			π-cation Lys 1239 x 3 π-sigma Lys 1239
22		Asn 1202 Lys 1045	π-cation Lys 1239 x 2 π-sigma Pro 1191

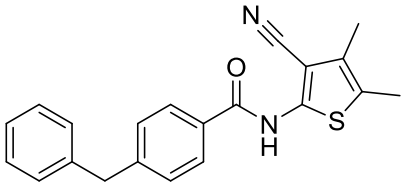
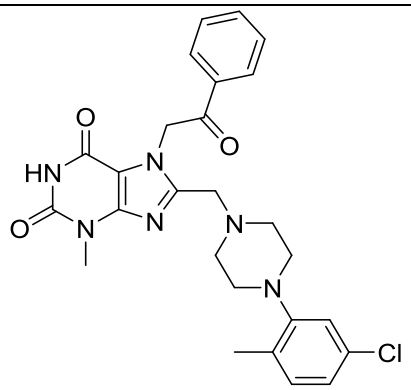
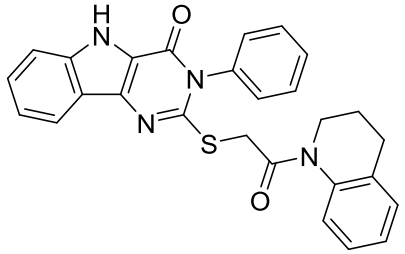
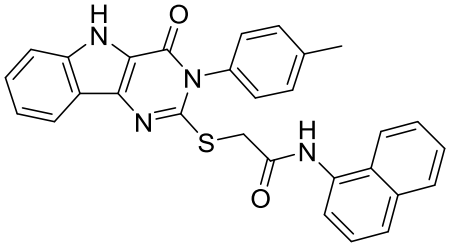
23		His 1222	π-cation Lys 1239 π-cation Lys 1045
24		Glu 1204 Asn 1202	π-sigma Pro 1191 π-sigma Lys 1239
25			π-cation Lys 1239 x 3 π-sigma Glu 1204
26			π-sigma Lys 1239 x 3

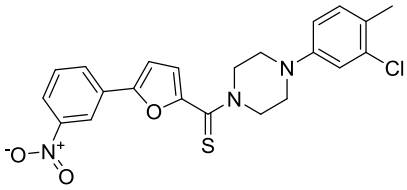
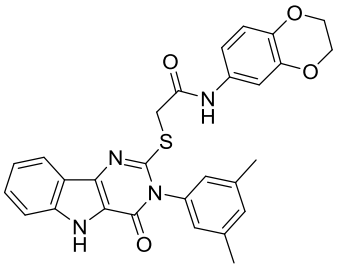
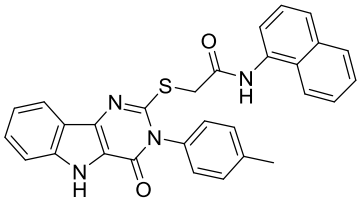
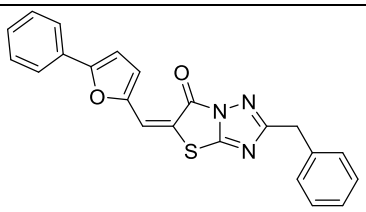
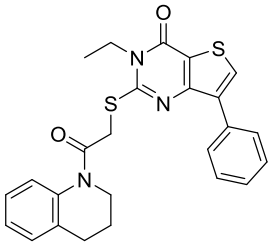
Table 5.2 Autodock output for the **nsP2_site_1** Fred hit list. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 1.0, 2.0 and 3.0 Å.

Entry*	Binding Energy *Kcal/mol*	Predicted K_i nM	Number of clusters&	Lowest energy cluster poses
2	-11.36	4.72	4	14
3	-7.26	4.79×10^3	8	13
4	-9.83	61.92	13	16
5	-7.22	5.12×10^3	9	7
6	-10.5	19.96	7	28
7	-9.46	116.7	7	22
8	-10.0	46.58	7	21
9	-8.68	435.31	17	20
10	-8.39	709.36	4	10
11	-9.88	57.32	7	16
12	-7.13	5.9×10^3	21	5
13	-10.16	35.58	1	50
14	-9.03	241.5	21	4
15	-7.69	2.29×10^3	5	28
16	-9.37	134.58	10	16
17	-8.45	644.95	4	22
18	-9.03	240.22	4	1
19	-9.18	185.76	2	48
20	-9.9	55.24	11	15
21	-9.75	71.74	20	5
22	-7.92	1.57×10^3	6	20
23	-8.3	822.33	2	44
24	-8.63	472.65	41	1
25	-9.89	56.66	7	12
26	-10.24	30.98	10	25

& Predicted docked poses are grouped together according to their root-mean-square (RMS) values compared to the input ligand structure. Poses that have similar RMS values (within the tolerance value selected) are grouped in one cluster.

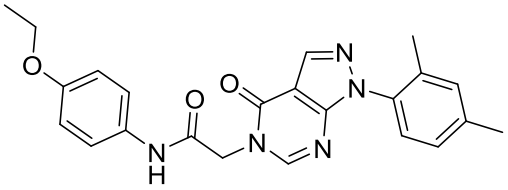
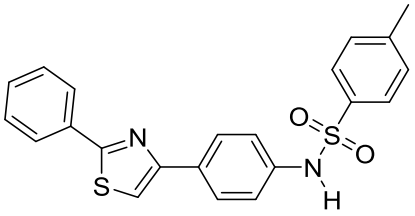
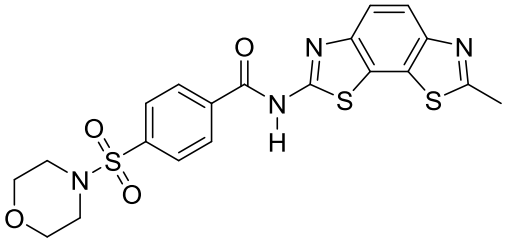
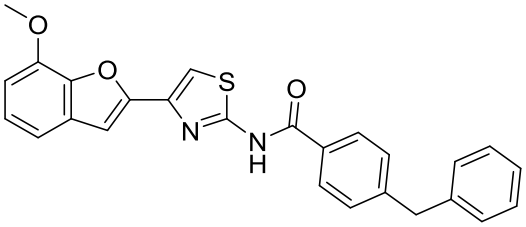
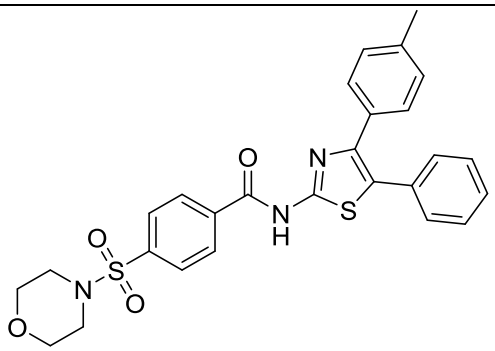
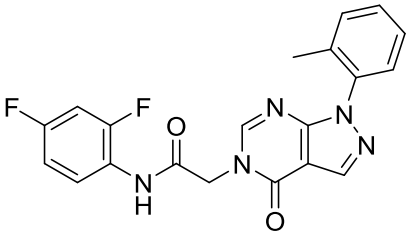
* Entries are the same as in Table 5.1, entry 1 was assigned for the Mg ion.

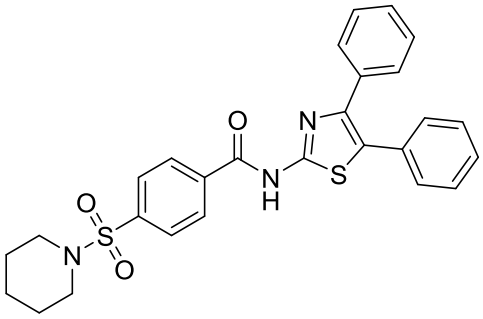
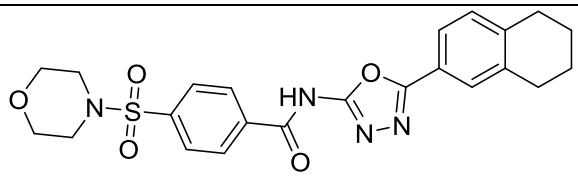
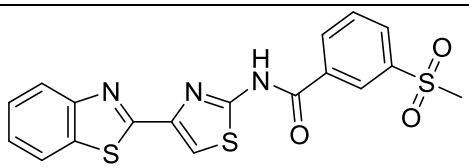
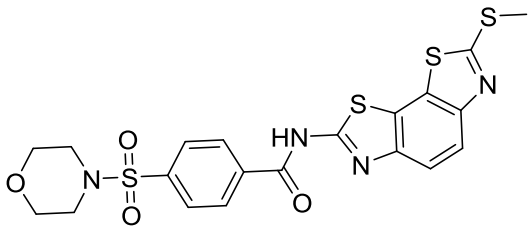
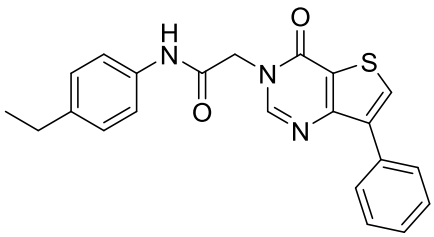
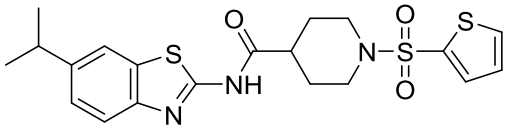
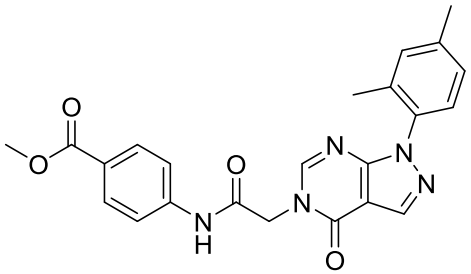
Table 5.3 The top 5 poses for the **nsP2_site_1** based on Autodock ranking, their ClogP, calculated binding energies, K_i values and interacting residues.

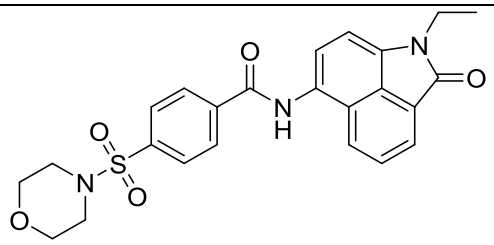
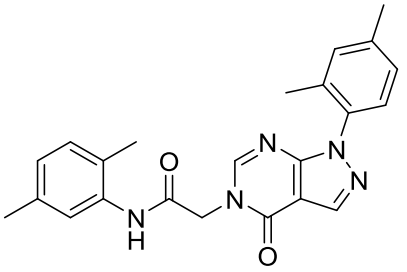
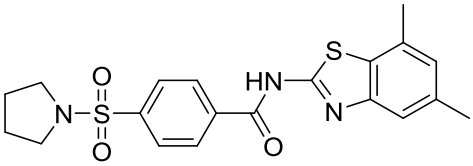
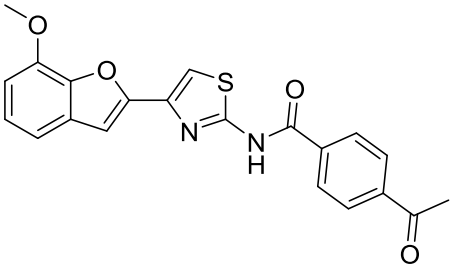
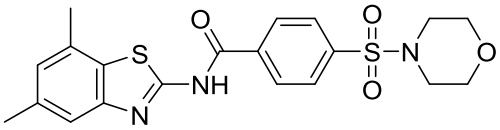
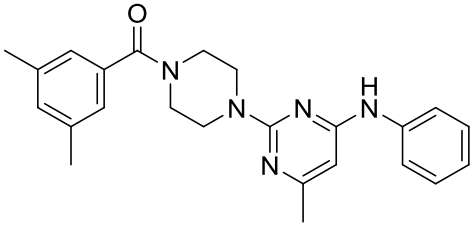
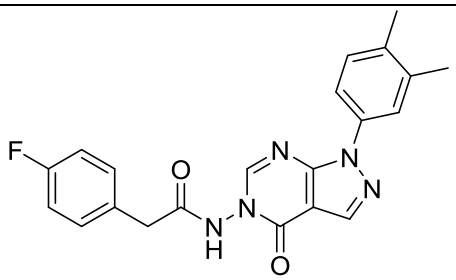
Entry	Structure	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues
2		4.08±0.69	-11.36	4.72	Lys 1045 (Hb, 1.9 Å) Lys 1239 (2Hb, 1.9, 2.5 Å) Lys 1239 (π-cation) Leu 1192 (π-sigma)
6		5.13±0.97	-10.50	19.96	Glu 1043 (2Hb, 2.1, 2.2 Å) Lys 1239 (Hb, 1.7 Å) Lys 1239 (π-cation) x 3
26		5.88±0.94	-10.24	30.98	Glu 1043 (Hb, 1.9 Å) Lys 1239 (Hb, 1.7 Å) Lys 1239 (π-cation) Lys 1239 (π-sigma)
13		6.07±0.88	-10.16	35.58	Lys 1045 (Hb, 2.4 Å) Lys 1239 (Hb, 2.1 Å) Pro 1191 (π-sigma) Lys 1045 (π-cation) Lys 1239 (π-cation)
8		5.67±1.05	-10.0	46.58	Lys 1045 (2Hb, 2.1, 2.2 Å) Lys 1239 (Hb, 1.8 Å) Lys 1045 (π-cation) Lys 1239 (π-sigma) x 2 Lys 1239 (π-cation)

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada).

Table 5.4 The top 25 compounds docked by Fred in **nsP2_site_2** and their interacting residues. Numbering is based on Fred output list (visualised by VIDA).

Entry	Structure	H-bonding residues	Other interacting residues
1		Trp 1084 Gln 1241	π-cation Lys 1091
2		Ser 1048	π-cation Lys 1091
3		Trp 1084 x 2 Asn 1202	
4		Ser 1048 Trp 1084	
5		Trp 1084 Asn 1202	
6		Trp 1084 Gln 1241	π-cation Lys 1091

7		Trp 1084	
8		Trp 1084 x 2 Asn 1202	
9		Trp 1084 x 2	π-sigma Ala 1046
10		Asn 1202 Trp 1084 x 2	
11		Trp 1084 Gln 1241	π-cation Lys 1091
12		Ser 1048 Trp 1084	
13		Trp 1084	π-cation Lys 1091

14		Trp 1084	π-sigma Mse* 1242
15		Trp 1084 Gln 1241	π-cation Lys 1091
16		Trp 1084 x 2	
17		Ser 1048 Trp 1084	
18		Trp 1084 Asn 1202	
19		Trp 1084	
20		Asn 1082	

* Selenomethionine

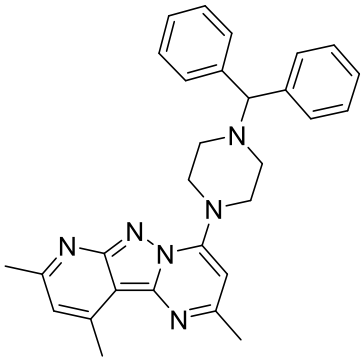
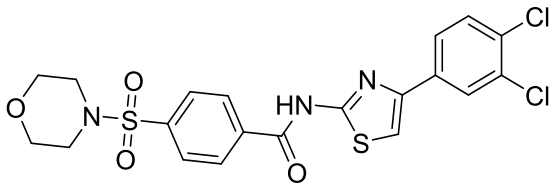
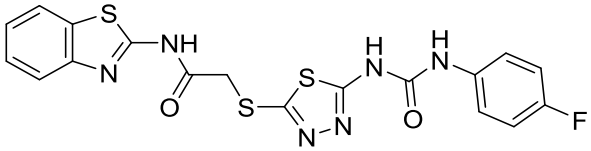
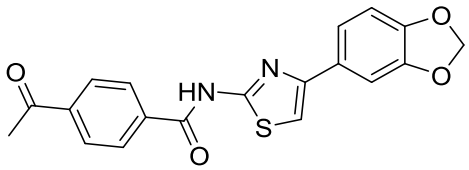
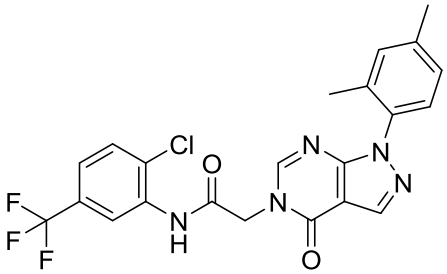
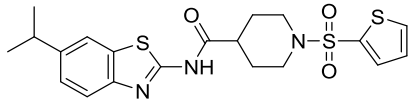
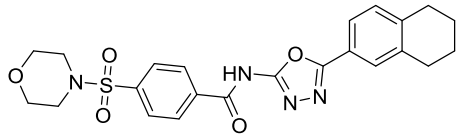
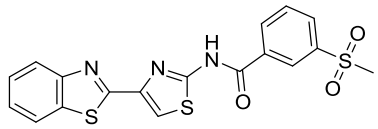
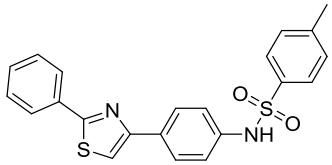
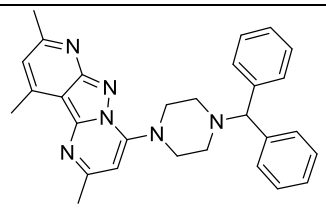
21		Ser 1048 Trp 1084	π-sigma Trp 1084
22		Asn 1202 Trp 1084	
23		Trp 1084 x 2 Ser 1048	π-cation Lys 1091 x 2
24		Trp 1084 Ser 1048	
25		Gln 1241 Tyr 1047	π-cation Lys 1091

Table 5.5 Autodock output for the **nsP2_site_2** Fred hit list. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 1.0, 2.0 and 3.0 Å.

Entry*	**Binding Energy *Kcal/mol***	**Predicted K_i nM**	**Number of clusters**	**Lowest energy cluster poses**
1	-8.32	799.94	13	7
2	-9.74	72.63	3	42
3	-9.24	167.47	3	31
4	-8.97	264.9	11	16
5	-8.8	353.41	5	2
6	-7.65	2.47×10^3	9	8
7	-9.23	170.77	6	10
8	-9.8	65.79	2	48
9	-9.76	70.47	4	30
10	-9.28	156.76	4	20
11	-8.42	667.9	7	16
12	-9.9	55.69	3	31
13	-8.38	722.18	9	20
14	-8.8	355.69	7	25
15	-8.88	312.02	10	14
16	-8.72	404.11	7	28
17	-8.45	636.79	7	25
18	-8.82	345.39	6	19
19	-9.4	129.44	7	27
20	-8.23	929.53	9	13
21	-9.67	81.91	1	50
22	-9.2	180.54	6	36
23	-8.68	431.68	9	2
24	-8.15	1.06×10^3	7	16
25	-8.53	562.66	12	5

* Entries are the same as in Table 5.4.

Table 5.6 The top 5 poses in the **nsP2_site_2** based on Autodock ranking, their ClogP, calculated binding energies, K_i values and interacting residues.

Entry	Structure	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues
12		4.57±0.65	-9.90	55.69	Ser 1048 (Hb, 1.6 Å) Tyr 1079 (Hb, 2.1 Å) Trp 1084 (Hb, 1.8 Å) Gln 1241 (Hb, 2.4 Å) Lys 1091 (π-cation)
8		4.13±0.68	-9.80	65.79	Trp 1084 (Hb, 1.7 Å) Lys 1091 (Hb, 1.8 Å)
9		3.28±1.00	-9.76	70.47	Glu 1050 (Hb, 2.3 Å) Tyr 1079 (Hb, 2.3 Å) Trp 1084 (2Hb, 2.1, 2.2 Å) Lys 1091 (Hb, 1.9 Å)
2		6.12±0.49	-9.74	72.63	Ser 1048 (Hb, 1.7 Å) Trp 1084 (Hb, 2.2 Å) Lys 1091 (π-cation)
21		4.86±1.49	-9.67	81.91	Trp 1084 (Hb, 1.8 Å) Lys 1091 (π-cation) x2

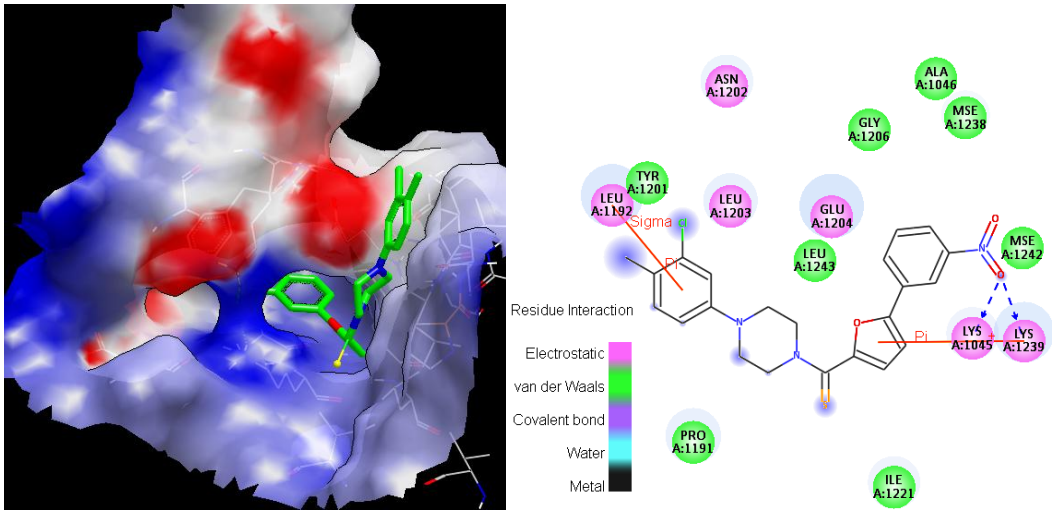
*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada).

5.3.1. *Analysis of the C domain virtual screening results*

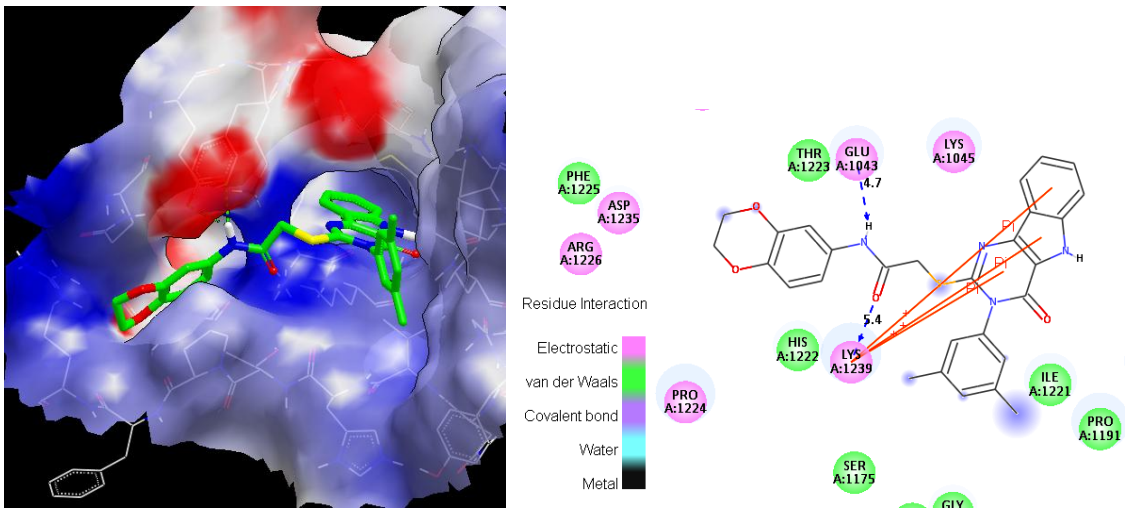
The CHIKV nsP2 plays an important role during the replication life cycle, and is composed of two domains. Domain C, which is the proteolytic machinery, processing the non-structural protein complex into individual functional proteins, and the N domain, which performs regulatory cellular functions through the NTPase and RTPase activities. These vital roles of the enzyme with the availability of a crystal structure, makes it an attractive target for the drug design for this virus. We investigated both domains for the possible binding sites. Within domain C which contains 3 cysteine residues (Figure 5.2) that might be involved in the proteolytic mechanism, access to these residues is most likely to occur through the site_1 (Figure 5.4), a possible site for

holding the protein to be cleaved by the protease activity. The site extends between domains N and C with residues from both lining the site_1 tunnel. The tunnel opens to the outer surface as a wide mouth lined mainly by residues from domain C. Our hypothesis is that bound small molecules to this site, not only freezes both domains together, preventing its function, but also blocks the sites through π-stacking interactions, mostly, hydrophobic/aromatic interactions. Our study did not reveal the participation of the cysteine residues (Cys1233, Cys1290), however, other residues have been found to strongly participate in the interactions in this site. Figure 5.7 shows the 3D and the 2D representations of the docked poses.

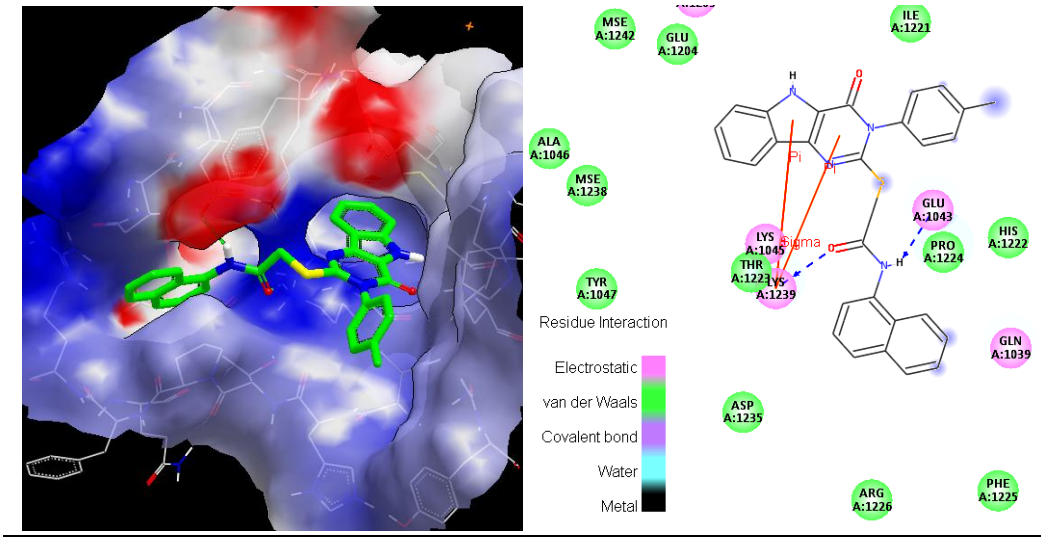
Pose 2



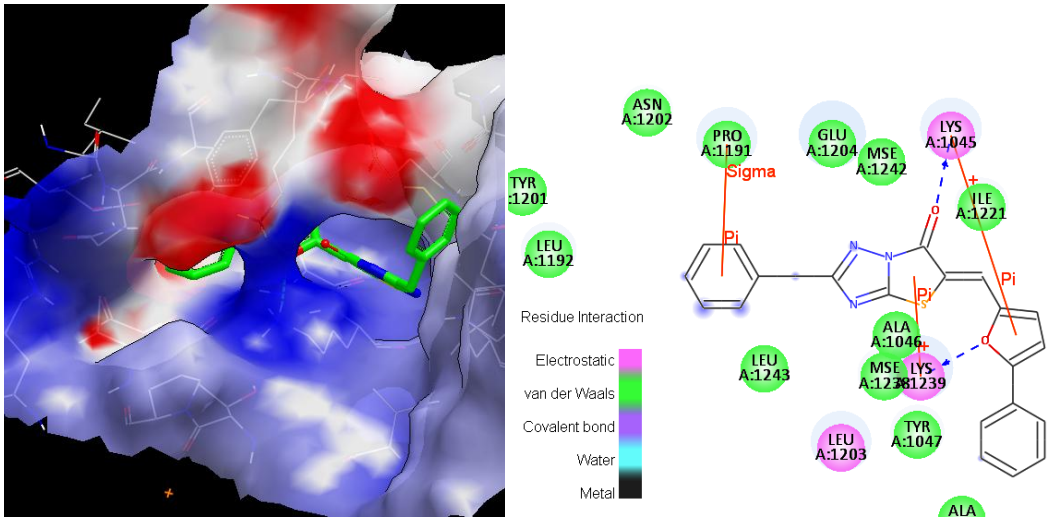
Pose 6



Pose 26



Pose 13



Pose 8

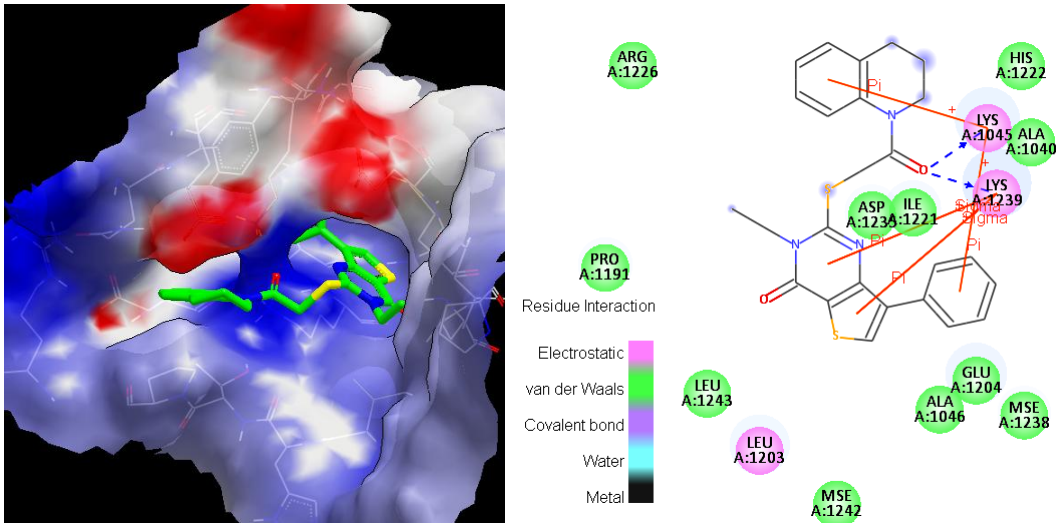


Figure 5.7 Three and two dimensional representations of the top 5 ranked poses (Table 5.3) inside the nsP2 site_1 pocket.

Figure 5.7 shows that compounds **2**, **13** and **8** were able to insert at least a phenyl ring inside the U-shaped tunnel whereas poses **6** and **26** were found to prefer the outer surface of the cavity (the outer mouth cavity). All compounds were able to bind to Lys1045 or Glu1043 from the N domain and to Lys1239 from the C domain. These interactions with residues in the two domains N and C make the compounds good possible ligands for this pocket. When bound to the pocket in the predicted way, compounds will be able to freeze domain N in relation to domain C, blocking the protease cavity of the enzyme, rendering the access to the cysteine residues for the catalytic cleavage not possible.

Poses **2** and **13** were able to form π stacking interactions with the side wall of the pocket *via* interactions with Leu1192 and Pro1191, respectively (Figure 5.7). This pocket side wall represents the enzyme outer surface, with this part of the enzyme surface being close to the linker cavity of site_2 (see Figure 5.6).

Poses **2** and **13** were able to orient in such way that they have aromatic rings relaxing before the side wall of the pocket forming interactions with Leu1192 and Pro1191, respectively. This unique ability for the two ligands (**2** and **13**) was related to: a) the presence of a relatively smaller linear terminal phenyl-furan sequence, such that this moiety can insert inside the tunnel, being stabilized by H-bonds with the tunnel hydrophilic bridge, b) the flexibility of the other terminal phenyl ring; in the docked pose **2**, the terminal phenyl is attached to the relatively flexible non-aromatic piperazine moiety whereas in the docked pose **13**, the terminal phenyl is attached to rigid aromatic central moiety *via* a flexible CH_2 linker. This flexibility enables the terminal phenyl rings to find the right position for interaction with the amino acid residues Leu1192 and Pro1191. These two classes of compounds can be described as relatively small rigid aromatic head groups linked in a linear manner through relative flexible joints to small aromatic tails. This orientation allows the insertion inside the tunnel forming stacking interactions with the cavity residues.

Poses **6** and **26** are derivatives of the main skeleton of *N*-aryl-2-((4-oxo-3-aryl-4,5-dihydro-3*H*-pyrimido[5,4-*b*]indol-2-yl)thio)acetamide, with different aryl groups. The 3,5-dihydro-4*H*-pyrimido[5,4-*b*]indol-4-one head is important in this class of compounds as it forms a strong π-stacking interaction with the C domain central Lys1239 (Figure 5.7), and is able to insert one the terminal phenyl rings inside the right

opening of the tunnel (on the right hand side when looking to front Figure). In both derivatives, the 2-mercaptoacetamide linker is important; the carbonyl oxygen can accept H-bond from the polar head of Lys1239 of the tunnel central bridge, and the NH participates in H-bonding with the N domain Glu1043. The aromatic tail; the relatively more flexible benzo[1,4]dioxane in compound **6** and the rigid naphthalene moiety in compound **26** blocked the access of the left opening of the tunnel (on the left hand side when front looking to the Figure). These three segments, bulky aromatic head, mercaptoacetamide linker and the aromatic tail were responsible for the orientation of this class in such a way that they block the entrance of the U-shaped tunnel from the two openings, and therefore, can be developed as selective blockers for this site. Figure 5.8 shows the superimposition of compound **6** and **26** inside site_1.

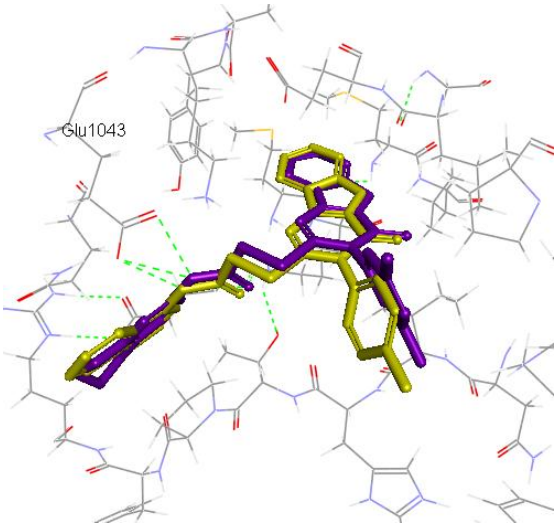


Figure 5.8 Superimposition of pose **6** (violet) and pose **26** (yellow) inside the nsP2_site1 binding pocket.

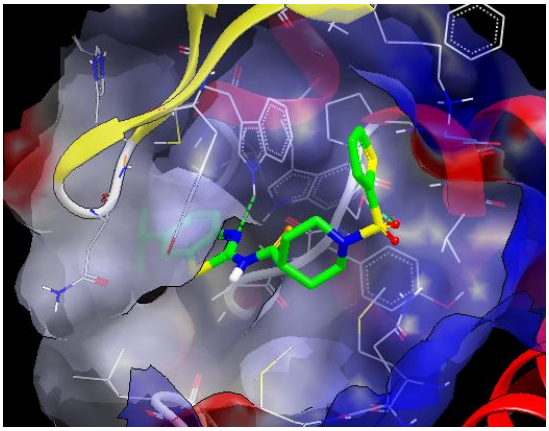
The mercaptoacetamide linker was also important in pose **8** where it connected the bulky head the phenylthieno[3,2-*d*]pyrimidin-4(3*H*)-one to the tail of the tetrahydroquinoline. The terminal phenyl substituent of the head was inserted inside the right hand side opening of the tunnel and was stabilized by interaction with the N domain tunnel bridge residue Lys1045 (Figure 5.7), whereas the main thieno[3,2-*d*]pyrimidin-4(3*H*)-one moiety was stabilized by interaction with the C domain tunnel bridge residue Lys1239 (Figure 5.7). The mercaptoacetamide linker forms H-bonds with both residues of the bridge tunnel residues. The benzo moiety of the tail forms an

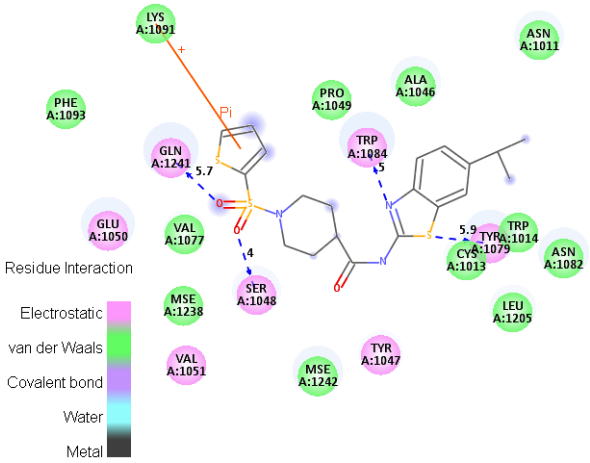
interaction with Lys1045 blocking the left tunnel opening. In comparison with poses **6** and **26** where the linker is attached to the rigid benzo moiety of the head groups, the linker in pose **8** is attached to the relatively flexible moiety of the rigid tail where certain flexibility might allow such stabilizing interaction with the tunnel bridge. Such stabilization was not possible in case of **6** and **26**.

5.3.2. *Analysis of the N domain virtual screening results*

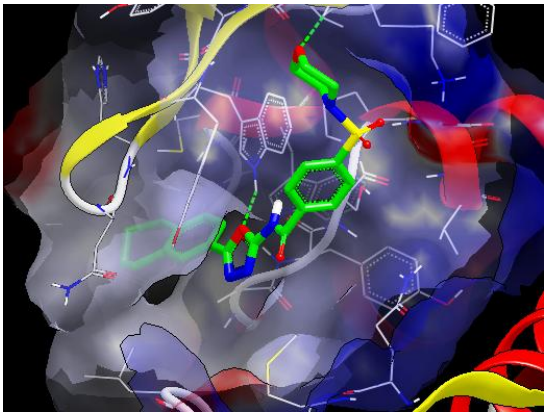
Site_2 is located within the N domain and lies as a channel that is partially covered by two of the N domain β-sheets. It represents the main pocket within domain N that might be important for the N domain functions. This pocket can also lead to access to the cysteine residues within C domain. Therefore, the hypothesis is that blocking the pocket will inhibit the enzymatic functions of this domain, and secondly, it might be an allosteric site for blocking the enzymatic activity of domain C. Figure 5.9 shows the 3D and 2D representations of the top ranked docked poses (Table 5.6) within the nsP2 site_2 N domain pocket.

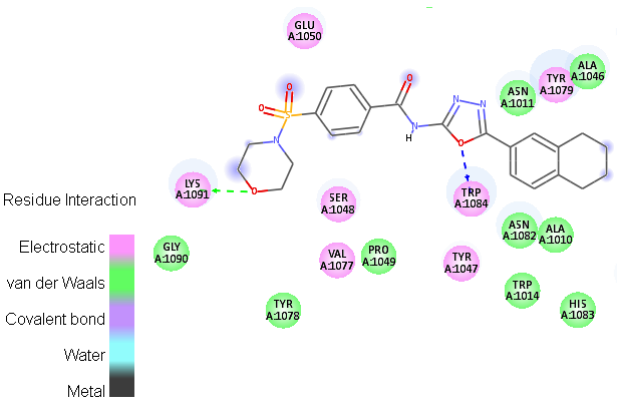
Pose 12



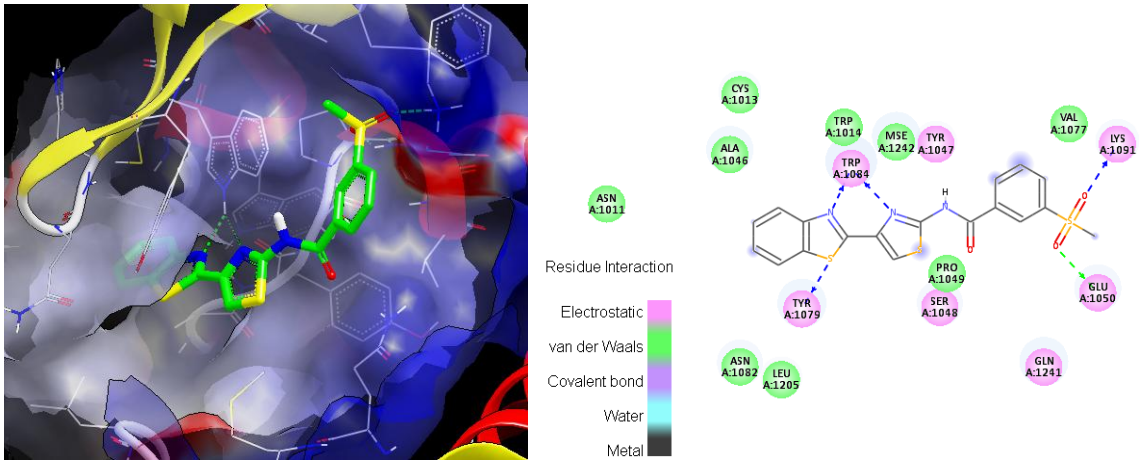


Pose 8

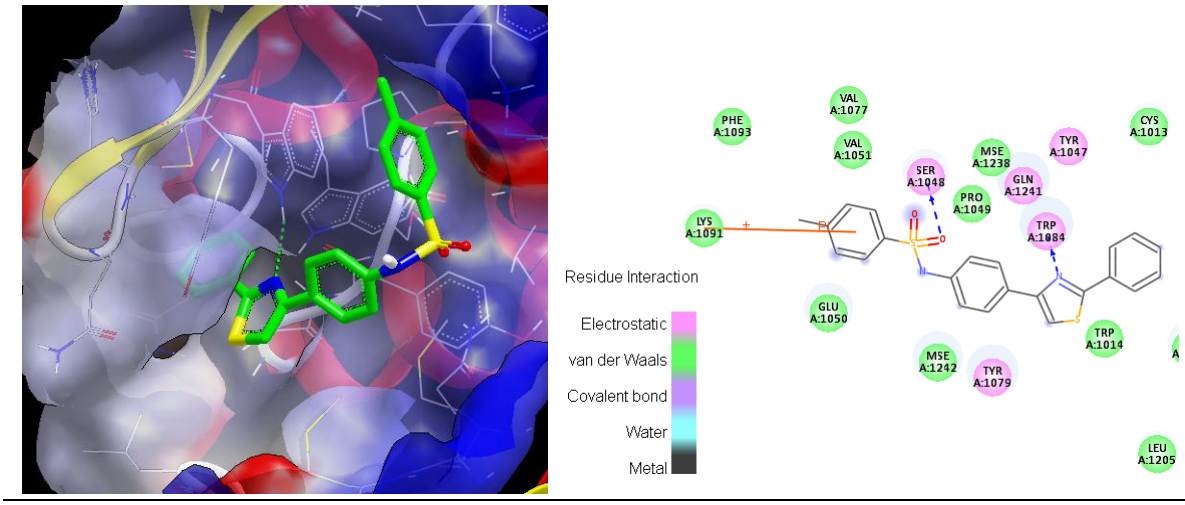




Pose 9



Pose 2



Pose 21

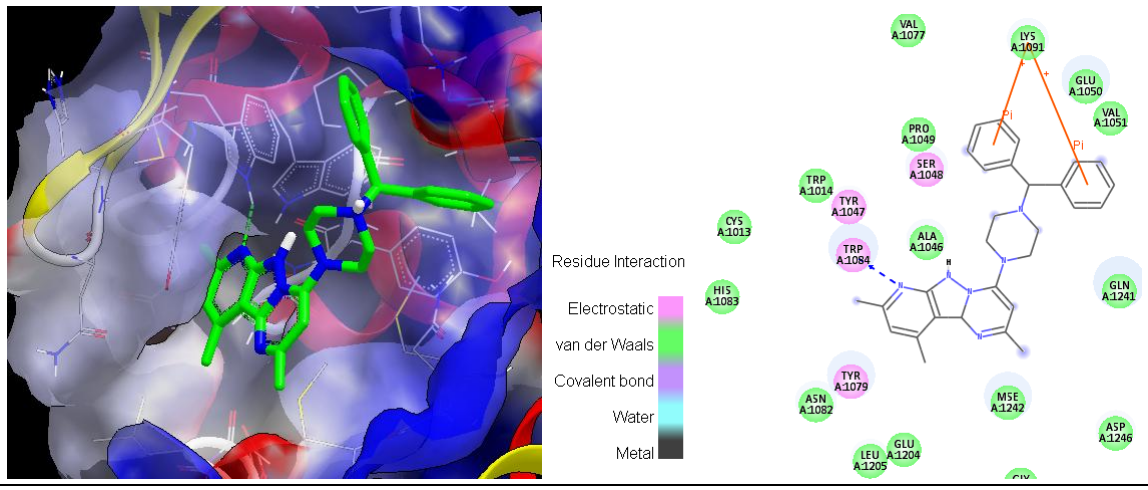


Figure 5.9 Three and two dimensional representations of the top 5 ranked compounds (Table 5.6) inside the nsP2 site_2 pocket.

Site_2 can be divided into 3 sections (Figure 5.5), the narrow linker cavity which opens towards site_1 side, a middle section and a terminal wide mouth. Two important interactions emerged in the top ranked poses, one strong H-bond with Trp1084 (in the range of 1.7-2.2 Å) which is located in the middle section and another interaction with Lys1091 in the terminal wide mouth section (Figure 5.9, the snapshots were taken with the wide mouth component in the front). This later interaction was either through H-bonds with the polar part of Lys1091 as in poses **8** and **9**, or through π-cation stacking interaction with the positively charged head of Lys1091 as in compounds **12**, **2** and **21**. In poses **12**, **8**, **9** and **2**, the terminal hydrophobic/aromatic moiety was inserted inside the narrow linker cavity of the pocket (Figure 5.9, the rear part of the snapshots).

All compounds interacted with residues within domain N without interacting with domain C except pose **12**, which forms a strong H-bond (2.4 Å, Figure 5.9) with the Gln1241 residue, located in one of the α-helices of domain C. This unique interaction was responsible for the better binding profile compared to the other docked poses as can be inferred from Table 5.6 where pose **12** was predicted to bind to the pocket with a binding energy of -9.90 kcal/mol with a predicted inhibition constant (K_i) of 55.69 nM.

The presence of the bi-heterocyclic system in the middle of the pocket, namely the benzo[*d*]thiazole moiety, in poses **12** and **9** was responsible for H-bond formation through the sulfur atom with the N domain β-sheet Tyr1079 residue; this β-sheet extends outwards like an umbrella over the pocket.

The presence of a sulfoxide group in the third wide mouth pocket section (front view of the snapshots taken in Figure 5.9) is favourable as both oxygens of this group can accept H-bonds from the N domain Ser1048 (around 1.6 Å, Table 5.6) as in case of poses **12** and **2**, or with the near by amino acid Glu1050 (2.3 Å) in case of pose **9**.

Poses **12**, **8** and **9** can insert hydrophobic tails within the linker cavity that opens on site_1 side. This indicates that this site may affect the other protease site somehow through this linker cavity. Figure 5.10 shows a 3D representation of the nsP2 surface with pose **12** (Table 5.3, within site_1) and pose **8** (Table 5.6, within site_2). It shows the possible clashes between the compounds in the linker cavity. It also shows the

possible allosteric effect of ligands in site_2 on the protease active site. This opens other possible investigations into using two inhibitors together as an effective inhibitory mechanism for that critical enzyme in the viral life cycle.

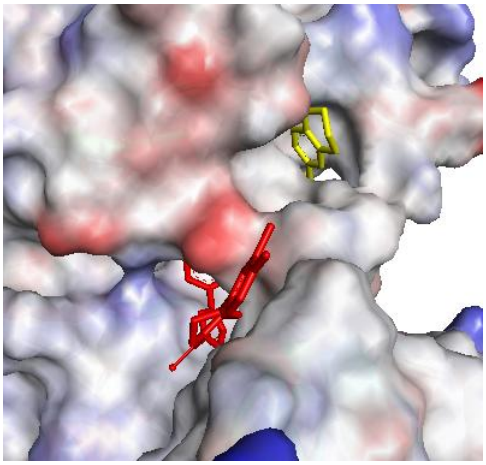


Figure 5.10 Pose **2** (Table 5.3) within the nsP2 site_1, shown in red colour. Pose **8** (Table 5.6) within the nsP2 site_2, shown in yellow colour. The hydrophobic tail of pose **8** within site_2 is protruding through the linker cavity towards pose **2** in site_1.

5.4. Concluding remarks

We have examined the CHIKV nsP2, a critical enzyme for the viral life cycle, for potential druggable pockets that could be targeted by small molecules. Two major pockets where identified within the N and C domains. The hypothesis described here investigated the possibility of inhibiting the enzyme multi-functions through blocking the active sites of the two domains. A database search on the two hydrophobic pockets was performed and this search detected the residues important for significant bindings as well as the preferred pharmacophores of the potential *in silico* inhibitors. Within domain C proteolytic site, Lys1239 and Lys1045 were found to interact with most of the identified ligands. Both residues represent a hydrophilic bridge in the C domain hydrophobic pocket that can form H bonds through their polar heads and stacking interactions through their side chains. Within domain N, Trp1084 and Lys1091 were found to be interacting with all of the docked poses, with the possibility for interaction with the C domain Gln1241. Some small molecules were also identified for each site that look promising *in silico* (nano-molar) inhibitors that might be starting points for

lead development targeting this enzyme. Another hypothesis is the possibility of using dual inhibitors for the enzyme, targeting both domains at the same time by a combination therapy. The CHIKV nsP2 enzyme is unfortunately not available for direct binding assays, therefore, the identified compounds were purchased and currently are being evaluated for their anti-chikungunya activities. This represents the first study in understanding the CHIKV nsP2 enzyme as well as identifying some selective potential inhibitors, and this computer model can be further used for screening large libraries to identify more hits. Depending on the activities, a retro-synthetic analysis can be designed for the active compounds, allowing in-depth *in silico* with *in vitro* optimization studies that might lead to the discovery of potent selective inhibitor for the virus.

CHAPTER 6: Structure Based Design towards the Identification of Novel Binding Sites and Inhibitors for the Chikungunya Virus Envelope Proteins

6.1. Introduction

The CHIKV genome is approximately 11.8 Kb in size and consists of a single stranded, positive sense RNA genome with two open reading frames (ORFs),[28] one in the 5` end which encodes two polyproteins, the precursors of the non-structural proteins. The second ORF at the 3` end encodes the structural proteins, the capsid (C), envelope glycoproteins E1 and E2 and two small cleavage products (E3, 6K). Like the other members of the *alphaviruses*, the CHIKV starts the life cycle by entering the target host cells by pH dependent endocytosis *via* a receptor mediated interaction.[30] A recent study identified prohibitin1 (PHB1) as a microglial cell expressed CHIKV binding protein.[31]

After entering the cell, the endosome acidic environment triggers conformational changes in the viral envelope complex made of E1 and E2 proteins, resulting in dissociation of the E2-E1 heterodimers and E1 homotrimers are formed. The E1 trimer inserts into the target cell membrane *via* its hydrophobic fusion peptide (fusion loop) and refolds to form a hairpin-like structure. Exposure of the E1 fusion peptide leads to releasing of the nucleocapsid into the host cell cytoplasm.[34,35] During the replication cycle inside the host cell, the capsid protein is released, and the pE2 and E1 glycoproteins are translated in the Golgi and are moved to the plasma membrane, where pE2 is cleaved by furin-like protease activity into E2 and E3.[42]

Glycoprotein E2 is responsible for receptor binding whereas E1 is responsible for membrane fusion.[10] E3 contains the 64-amino-terminal residues of p62 and mediates proper folding of pE2 and its subsequent association with E1.[93] E3 also protects the E2-E1 heterodimer from premature fusion with cellular membranes.[94] Furin maturation of p62 into E3 and E2 during transport to the cell surface primes the spikes for subsequent fusogenic activation for cell entry. Mature virions bud at the plasma membrane *via* interactions between E2 and genome-containing viral nucleocapsids present in the cytoplasm,[95] ready for infecting new cells. The crystal structures of both the immature and the mature glycoprotein complexes have recently been solved,[95] (Figure 6.1).

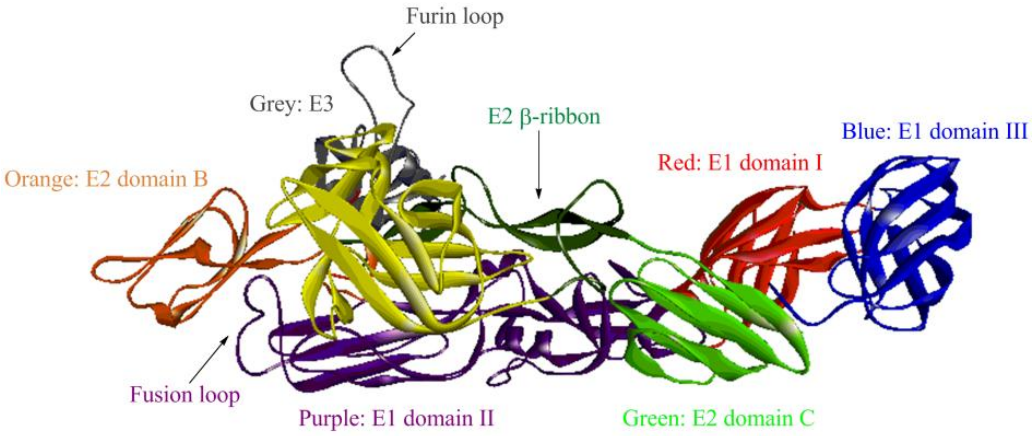


Figure 6.1 Crystal structure of the immature envelope glycoprotein complex of Chikungunya. Generated from the PDB file: 3N40.[95]

E1 is folded into three β-sheet rich domains (I, II and III). E2 is an all β protein belonging to the immunoglobulin superfamily, with three domains A, B and C. Domain B is at the membrane upper end and domain C is towards the viral membrane: Domain A is at the centre while domain C binds to the adjacent domain II of E1. The long β – ribbon of E2 makes most of the connection with E3. Furin loop is E2-E3 junction in the immature complex; this junction contains a functional proprotein convertase motif which is cleaved by the cellular porteases; furin-like proprotein convertases during the maturation of the glycoproteins.[42] The amino acid His60 in this junction is the critical residue that determines the spectrum of furin and furin-like convertases that process E2-E3 glycoprotein complex.[100] The U shaped fusion loop of E1 is inserted in a groove between E2 domains A and B being stabilized by hydrogen bonds with E2 histidine side chains.[95] In the neutral pH, E3 maintains E2 domain B in an orientation with respect to domain A in such way that it creates the groove accommodating the E1 fusion loop, protecting the virus from premature fusion with other cellular membranes.[95,101] Some residues in domain B of E2 are believed to be associated with cell recognition.[94] The fusogenic activity of the E1 fusion peptide is highly dependent on pH change. The histidine residues of E2 are believed to be involved as the pH sensor for the activation of the fusion protein at lower pH[95] due to the increased probability of histidines to become positively charged at lower pH values, based on the fact that the imidazole ring of the histidine residue is the only amino acid side chain whose apparent dissociation constant from protons (pKa) falls within the physiological range. Within the E1 fusion

peptide sequence, the glycine residue (Gly91) is critical for the fusion process. Also, it was found that one histidine residue at E1 230, which is located outside of the fusion sequence, is also critical for the fusion.[96]

Blocking the *in vitro* CHIKV infection in the host cells targeting the envelope proteins has been demonstrated by blocking the intracellular furin-mediated cleavage of viral envelope glycoproteins (E2E3 or p62 precursors), this blocking was achieved by an irreversible furin-inhibiting peptide which significantly reduce the processing of E3E2 CHIKV glycoproteins. This led to the formation of immature viral particles and impaired viral spreading through other uninfected cells.[100] This reflects the importance of considering the envelope glycoproteins as an attractive target for selective drug development.

The usage of the three-dimensional structure of the target proteins (crystal structures) in the virtual screening (*in silico*) of chemical libraries has been a powerful approach to identify lead compounds with some successful examples in a number of systems.[226, 227] Such structure based drug design techniques, including the identification of new binding sites and virtual screening search, have been successfully used for identification of lead compounds for the dengue virus envelope protein (E protein).[228, 229] Dengue virus is also an arbovirus and is transmitted by the same vector mosquito of the CHIKV. Herein, we report for the first time the novel binding sites in the CHIKV envelope glycoproteins that can be used as sites for inhibitors that could alter the function of the envelope proteins and consequently, inhibit the virus fusion function. To increase the chances of possible hits, we examined both the immature and the mature glycoprotein crystal structures for possible binding sites. Two sites were chosen that were common in both the immature and the mature proteins based on their locations and functions. We then used virtual screening combining two different docking algorithms with a number of chemical databases to identify suitable compounds predicted to bind in these sites. FRED (fast rigid exhaustive docking) was used for fast and precise screening using multiple scoring functions, followed by re-docking ranking of the top hits using AUTODOCK scoring function. This led to the identification of favoured hits that have suitable binding profiles to the CHIKV glycoproteins. This hypothesis represents a new strategy for inhibiting this particular virus by targeting the envelope proteins which will lead to impaired protein function and thus inhibiting the virus, and will help the further synthetic development and optimization of selective inhibitors, as

previously and successfully achieved for the dengue virus envelope protein inhibitors.[229, 230]

6.2. Results and discussion

6.2.1. Identification of novel binding sites

Both the crystal structure of the immature complex (PDB file: 3N40 [95]) and the mature complex (PDB file: 3N42[95]) were used. Binding sites within the receptors were detected using the Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012). The algorithm is based on a grid search and "eraser" algorithm which derives binding sites from cavities in the structure of the receptor. The binding site found is displayed as a set of points. The volume of each cavity is defined as the product of the number of site points and the cube of the grid spacing. Six main sites were detected in both the immature and the mature crystal structures and only one site was detected in the mature crystal structure that is not present in the immature form (Figure 6.2). Table 6.1 shows the identified sites. Suitable cavities were then checked further based on functionality, presence of hydrophobic residues, presence of charged residues and solvent accessibility.

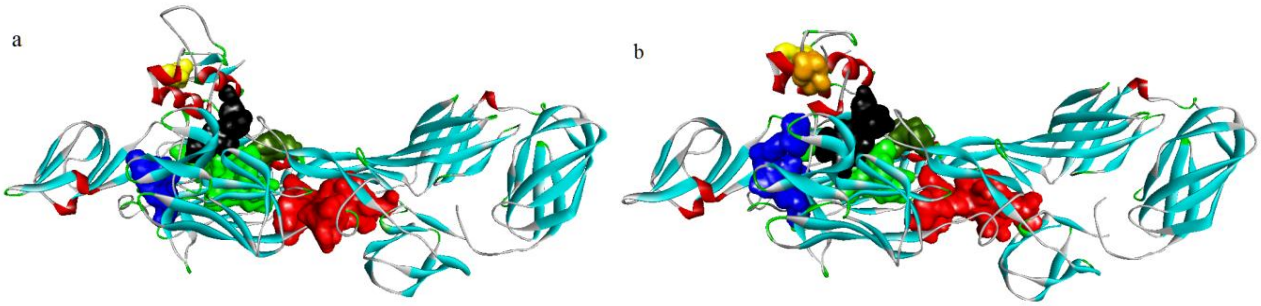


Figure 6.2 Crystal structure of (a) the immature glycoproteins (generated from file pdb: 3N40) and (b) the mature glycoproteins (generated from the file pdb: 3N42) showing the identified binding cavities as solid filled surface.

Table 6.1 The identified receptor cavities in the immature (3N40) and mature (3N42) crystal structures, grid coordinates x, y and z, cavity volumes, points count and location for each cavity site.

		3N40 (immature structure)	**3N42 (mature structure)**	**Location**
Site 1 (red)	x, y, z	-15.381, -1.269, 16.434	-15.687, 2.019, -19.939	Between E1 domain II and E2 domain C
	Volume	687.25	651.375	
	Points count	5498	5211	
Site 2 (light green)	x, y, z	-30.631, 17.481, 33.684	-33.937, -18.731, -31.939	Between E1 domain II and the β–ribbon of E2
	Volume	395	357.375	
	Points count	3160	2859	
Site 3 (dark green)	x, y, z	-30.631, 4.981, 37.934	-33.437, -6.731, -33.189	Adjacent to site 2
	Volume	157.625	156.125	
	Points count	1261	1249	
Site 4 (blue)	x, y, z	-38.131, 31.481, 24.934	-42.937, -28.731, -22.939	Behind the fusion loop, between E3, E2 domain B, E2 domain A
	Volume	126.25	183.875	
	Points count	1010	1471	
Site 5 (black)	x, y, z	-44.631, 14.731, 23.184	-44.437, -14.731, -23.439	between the β–ribbon of E2 and E3
	Volume	93.125	124	
	Points count	745	992	
Site 6 (yellow)	x, y, z	-57.631, 16.731, 36.184	-16.187, -18.231, -36.439	Within E3 cavity
	Volume	29.5	20.5	
	Points count	236	164	
Site 7 (orange)	x, y, z	Does not exist	-59.187, -15.731, -26.189	Replacing the furin loop
	Volume		22.25	
	Points count		178	

6.2.2. Virtual screening with the CHIKV envelope proteins

Two chemical compounds libraries were used; The NCI set library of 265,242 compounds and the Life Chemicals protein-protein interactions inhibitors library of 31,143 compounds. The databases were filtered with the drug-likeness-index; limit the

range for Molecular Weight ≤ 500, calculated octanol–water partition coefficient (clogP ≤ 5), and hydrogen bond donors, and acceptors (OH's and NH's ≤ 5; N's and O's ≤ 10),[231] using Filter v2.0.2 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com), producing 55,841 compounds from the NCI library and 4,124 compounds from the Life Chemicals library. Fast exhaustive virtual screening was performed using FRED v2.2.5 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com), which is a fast and effective docking application whose performance is significantly more reliable, i.e. lower variance, than most other programs.[232, 233] FRED performs a systematic, exhaustive, nonstochastic examination of all possible poses within the protein active site, filters for shape complementarily [234] and pharmacophoric features before selecting and optimizing poses using the Chemgauss scoring function. Omega2 (Systematic high-throughput conformer generation, OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com),[235, 236] was used to generate multiple conformers for each compound in the database libraries using the default settings. Omega2 takes into account the flexibility of a molecule by generating all representative conformers. For the NCI library, 2,312,012 conformers were generated, and 334,064 conformers were generated from the Life Chemicals compounds. The work-flow diagram is shown in (Figure 6.3), the life chemical library was screened on site 2 (light green colour in Figure 6.2) in both of the immature and the mature glycoproteins. The NCI set compounds were screened on site 4 (blue colour in Figure 6.2) of the two envelope protein forms. The binding sites were prepared for docking using Fred receptor setup software (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com). The grid boxes were determined based on the x, y and z co-ordinates given in Table 1. After the docking calculations, the poses returned were scored and ranked with a Gaussian shape function independently by the five available scoring functions (PLP, Chemgauss3, Chemscore, OEChemscore, and Screenscore) and by a consensus of all. The top ranked poses from the exhaustive docking were then optimized using systematic solid body optimization by chemgauss3. VIDA v4.2.0 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com) was used to visualise the docked poses within the receptor active site, and to inspect the critical interacting residues in each pocket with the individual docked poses. Top 20 hits were then recorded for each of the four sites (Tables 6.2, 6.5, 6.8, 6.11).

The top 20 docked poses ranked in each of the four binding sites were then extracted as PDB files, and were processed with Autodock Tools 1.5.6rc3 (ADT) graphical interface.[173] The Gasteiger charges were calculated and the nonpolar hydrogen atoms were merged, torsion angles were defined, they were then saved as pdbqt files for Autodock calculations. Crystal structures (3N40, 3N42) were used by Autodock Tools 1.5.6rc3 to setup the receptor binding sites. The grid box co-ordinates in each site were determined based on the co-ordinates in Table 1. AutoGrid 4.2 algorithm was used to evaluate the binding energies between the inhibitors and the enzyme and to generate the energy maps for the docking run. Fifty runs were generated using Autodock 4.2 Lamarckian genetic algorithm[173] for the searches. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0 Å, the docked poses were ranked according to the binding energies and ligand efficiencies, and finally the five lowest energy poses (Tables 6.3, 6.4, 6.6, 6.7, 6.9, 6.10, 6.12, 6.13) were selected as the resultant complexes with the enzymes. Compounds are commercially available and have drug like qualities and also can be accessed through chemical syntheses for further optimization process.

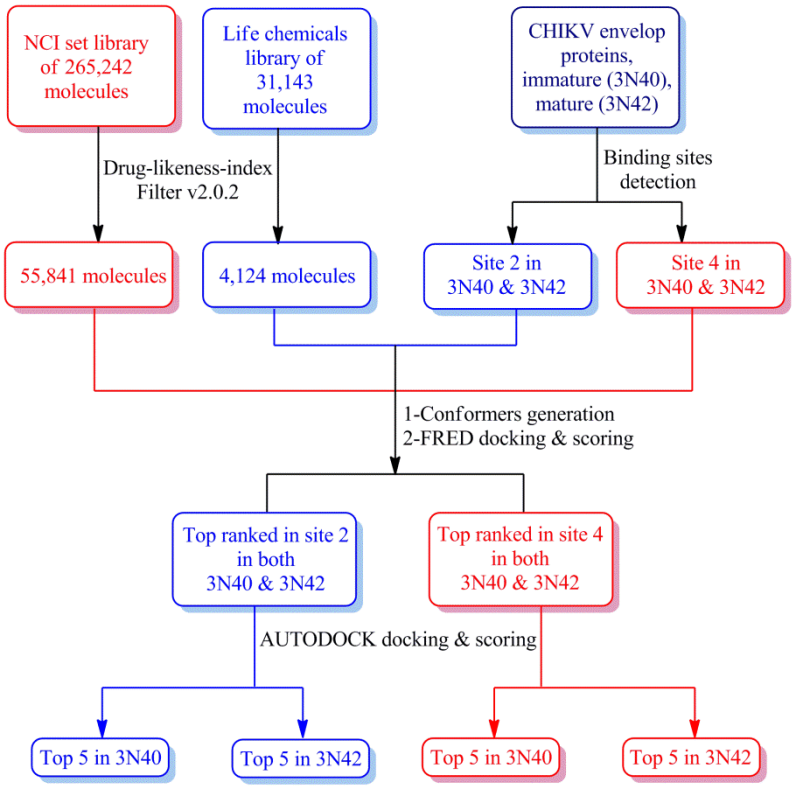
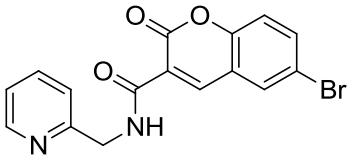
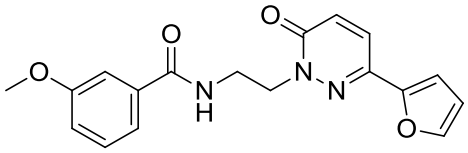
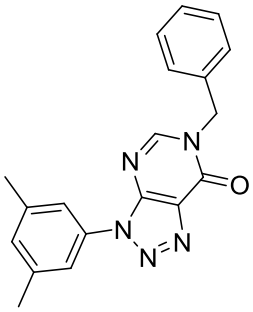
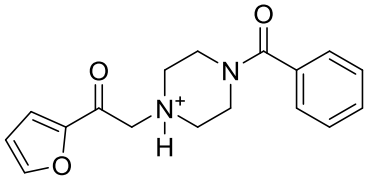
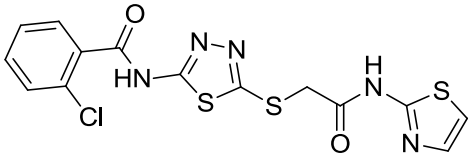
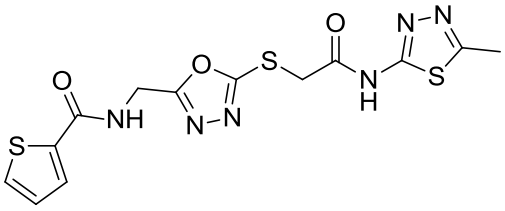
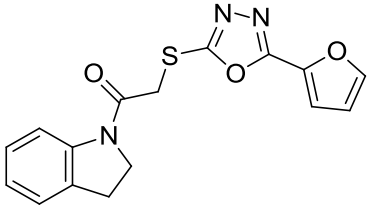
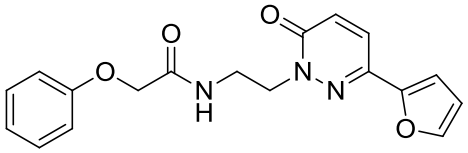
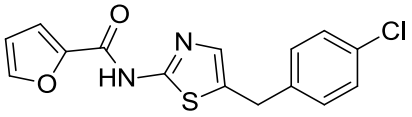
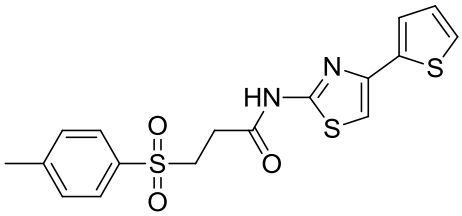
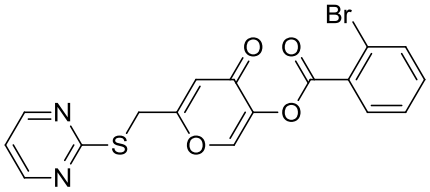
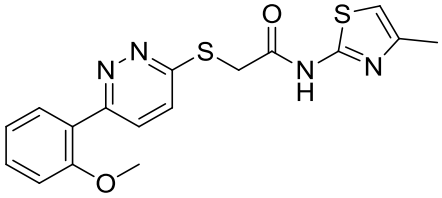
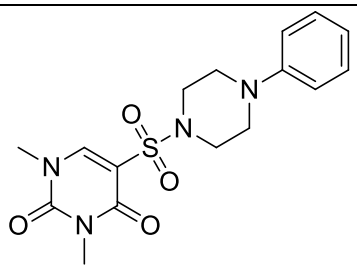
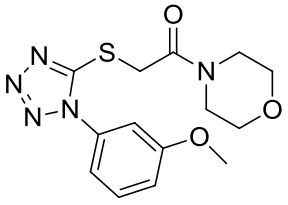
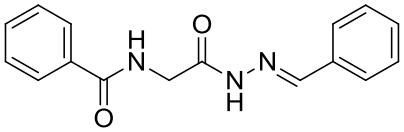


Figure 6.3 Work-flow diagram of the virtual screening procedures used for CHIKV envelope proteins.

Table 6.2 Fred output virtual screening top 20 hits and their target residues, for the *immature glycoprotein receptor* (3N40) site_2.

Entry	Structure	H-bonding	Other interactions
1		E1 Lys 52 E1 Thr 53 E2 Tyr 301	
2			π–sigma (E1 Val 54)
3			π–sigma (E2 Arg 100)
4		E1 Thr 53 E1 Ile 55 Lys 52	π–cation (E1 Lys 52)
5		E1 Ile 55	π–sigma (E1 Val 54)
6		E1 Val 231 E1 His 230 E1 Ile 55	π–cation (E1 Lys 52)
7		E1 His 230 E1 Val 231	

8		E2 Tyr 301	π–cation (E2 Arg 100)
9		E1 Ile 55 (2 Bonds) E2 Tyr 301 (2 Bonds)	
10			
11		E1 Thr 53, Ile 55 E2 Arg 100, Tyr301	
12			π–cation (E1 Lys52)
13		E2 Glu 232	
14		E1 Val 231, His 230	
15		E1 Tyr 301 E2 Glu 232	π–sigma (E1 Val 54)

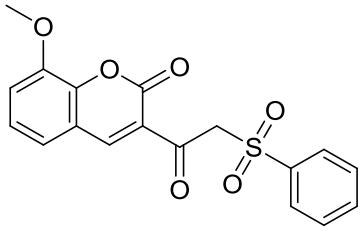
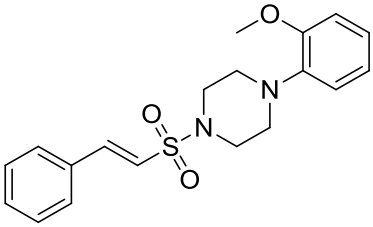
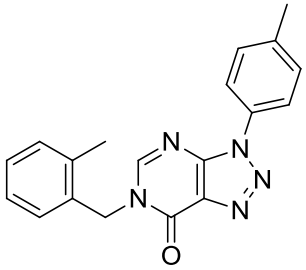
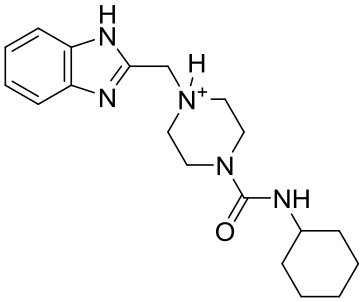
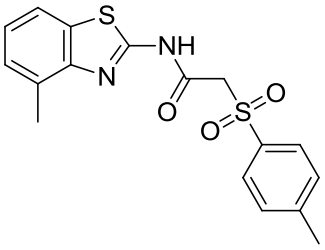
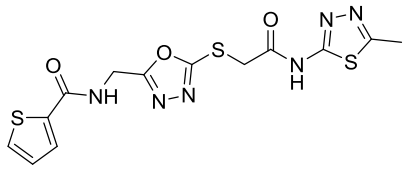
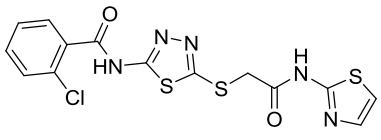
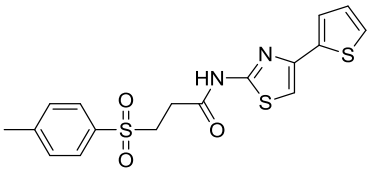
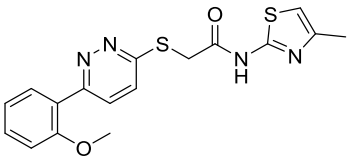
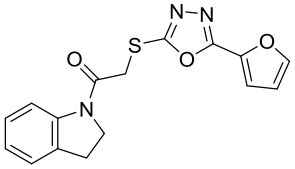
16		E1 Ile 55 x 2	
17			
18			
19		E1 Val 54, Thr 53 E2 Tyr 301	
20		E2 Tyr 301 x 2	

Table 6.3 Autodock output for the *immature glycoprotein receptor* (3N40) site_2 Fred hit list, Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0, 3.0 Å.

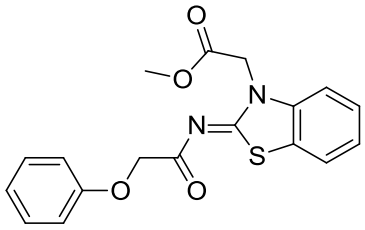
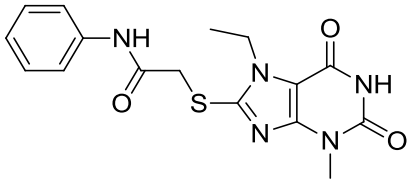
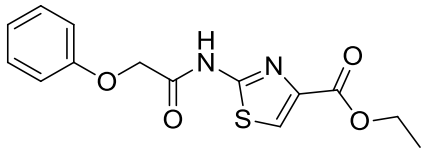
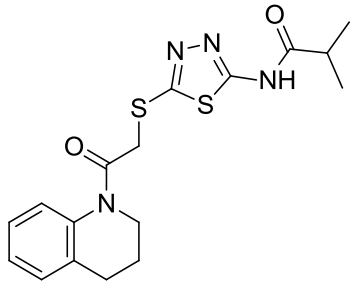
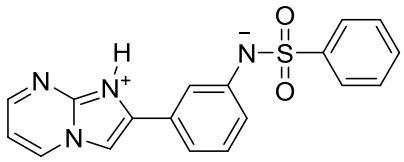
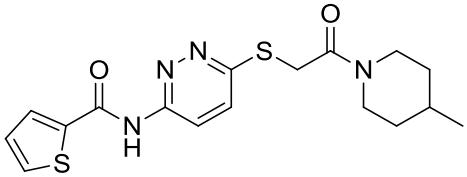
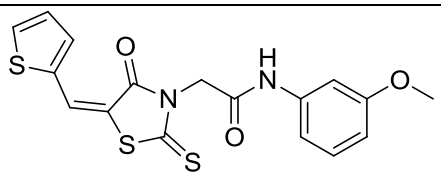
Entry	Binding Energy *Kcal/mol*	Predicted K_i (nM)	Number of clusters	Lowest energy cluster poses
1	-8.37	734.41	6	24
2	-8.23	930.55	12	19
3	-8.38	721.49	2	41
4	-7.67	2.39×10^3	5	24
5	-9.43	121.87	10	10
6	-10.06	42.15	10	15
7	-8.99	255.48	8	19
8	-8.04	1.29×10^3	13	9
9	-8.3	819.29	6	31
10	-9.36	138.19	6	9
11	-8.8	354.66	16	3
12	-9.18	187.59	14	3
13	-8.19	987.11	1	50
14	-7.91	1.6×10^3	16	3
15	-7.85	1.75×10^3	4	23
16	-8.8	357.02	7	11
17	-8.7	416.95	7	18
18	-8.43	657.11	4	22
19	-8.97	267.05	5	33
20	-8.98	263.45	9	13

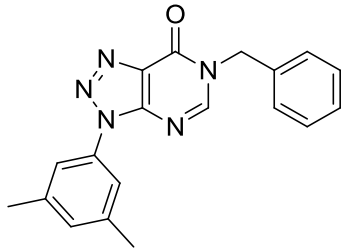
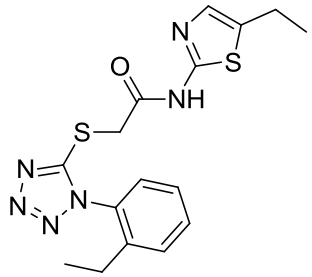
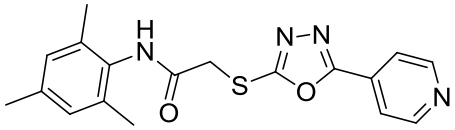
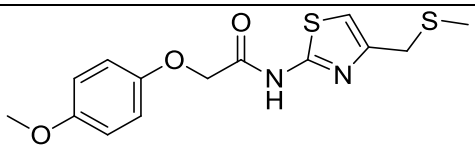
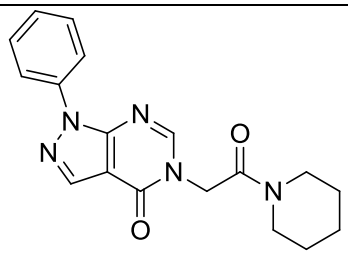
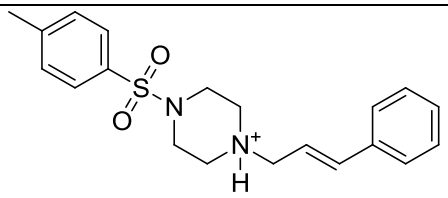
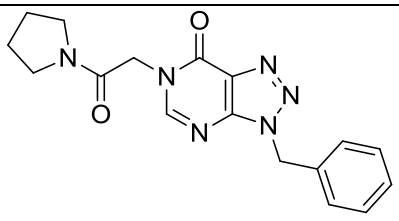
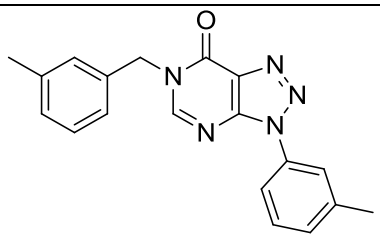
Table 6.4 Top 5 hits identified for site 2 using the *immature glycoprotein receptor* (3N40), showing the molecular weights, calculated logP (clogP), predicted binding energies, inhibitory constants (K_i) and the interaction residues.

	Structure	Molecular weight	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues
1		396.47	0.31±0.89	-10.06	42.15	E1 Lys52 E1 Ile55 E1 Thr53 E2 Tyr301 E2 Arg100
2		411.91	3.43±0.66	-9.43	121.87	E1 Lys52 E1 Ile55 E1 Thr53 E2 Tyr301 E2 Arg100
3		392.52	3.60±0.43	-9.36	138.19	E1 Lys52 E1 Ile55 E2 Tyr301 E2 Glu232
4		372.46	2.93±0.44	-9.18	187.59	E1 Lys52 E1 Ile55 E2 Tyr301 E2 Glu232 E2 Arg100
5		327.36	3.47±0.71	-8.99	255.48	E1 Lys52 E1 Ile55 E2 Tyr301 E2 Arg100

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada).

Table 6.5 Fred output virtual screening top 20 hits and their target residues, for the *mature glycoprotein receptor* (3N42) site_2.

Entry	Structure	H-bonding	Other interactions
1		E1 Ile 55, Lys 52 E2 Tyr 237, Arg 36	π–cation (E2 Arg 36) π–cation (E1 Lys 52)
2		E1 Ile 55, Lys 52, Glu112 E2 Tyr 237, Glu 66	π–cation (E1 Lys 55) π–cation (E2 Arg 36) π–sigma (E2 Ile 167)
3		E1 LYS 52 x 2	π–sigma (E1 Val 54)
4		E1 Lys 52 E2 Tyr 237	
5		E1 Ile 55, Lys 52 E2 Glu 35, Tyr 237	π–cation (E1 Lys 52)
6		E1 Lys 52 E2 Glu 35	
7		E1 Ile 55	π–sigma (E1 Val 54) π–cation (E1 Lys 52)

8		E1 Lys 52	π–sigma (E2 Arg 36) π–cation (E1 Lys 52)
9		E1 Ile 55 E2 Tyr 237 x 2	π–cation (E1 Lys 52) π–cation (E2 Arg 36)
10		E1 Ile 55, Val 231 E2 Tyr 237	π–cation (E1 Lys52)
11		E1 Ile 55, Lys 52 x 2	π–sigma (E1 Val 54)
12		E1 Lys 52	π–sigma (E1 Val 54)
13		E2 Tyr 237	π–cation (E1 Lys 52)
14		E1 Ile 55	
15		E1 Lys 52	π–cation (E2 Arg 36)

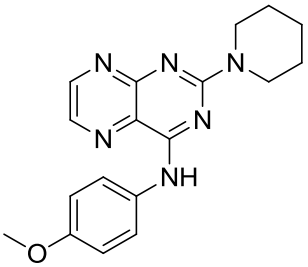
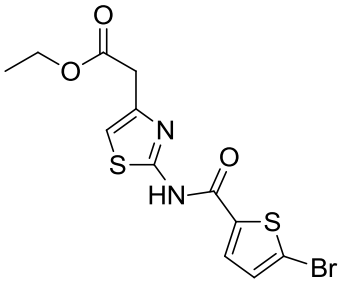
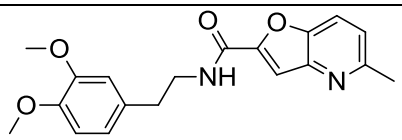
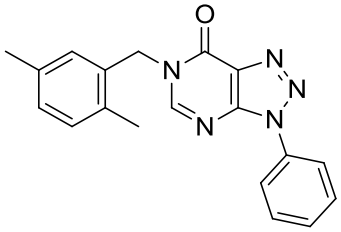
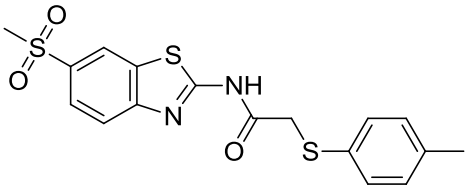
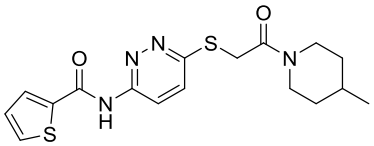
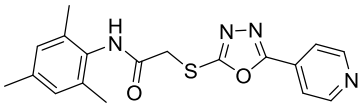
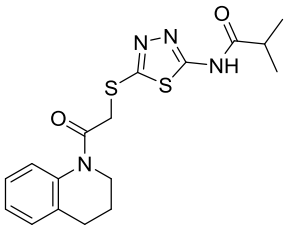
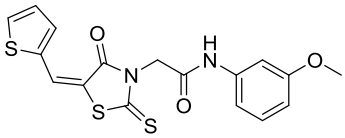
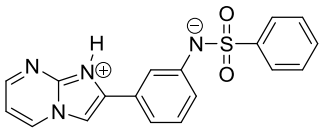
16		E1 Lys 52	π–cation (E2 Arg 36)
17		E1 Lys 52 x 2	π–sigma (E1 Val 54)
18		E1 Lys 52 x 2	π–cation (E2 Arg 36) π–cation (E1 Lys 52)
19		E1 Lys 52	
20		E1 Lys 52	π–cation (E1 Lys 52) π–cation (E2 Arg 36)

Table 6.6 Autodock output for the *mature glycoprotein receptor* (3N42) site_2 Fred hit list, Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0, 3.0 Å.

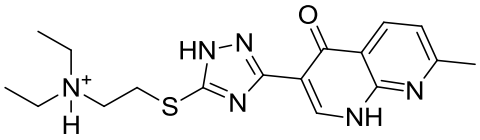
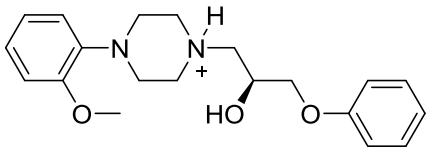
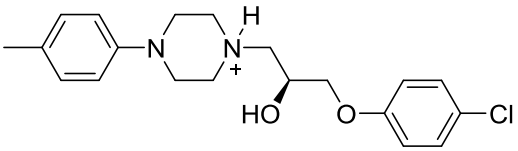
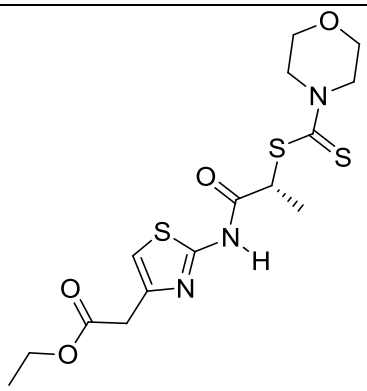
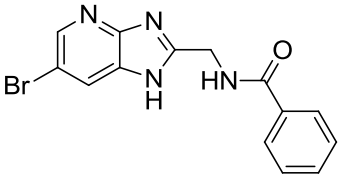
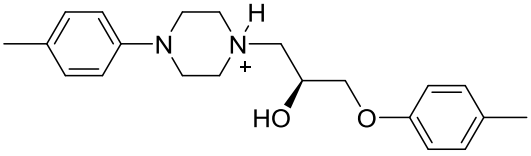
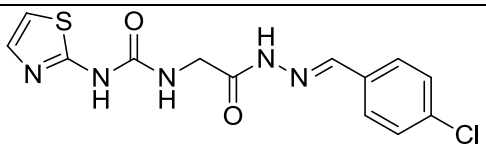
Entry	Binding Energy Kcal/mol	Predicted K_i (nM)	Number of clusters	Lowest energy cluster poses
1	-8.8	357.4	8	15
2	-8.51	581.05	10	23
3	-7.63	2.57×10^3	14	11
4	-9.36	138.07	8	30
5	-9.17	190.7	18	2
6	-9.98	48.38	8	22
7	-9.26	163.4	7	23
8	-8.05	1.25×10^3	5	13
9	-8.98	262.77	16	15
10	-9.71	75.78	8	14
11	-7.83	1.81×10^3	10	12
12	-7.97	1.44×10^3	9	28
13	-8.65	454.1	8	2
14	-8.29	842.46	8	9
15	-8.53	554.46	5	22
16	-8.36	741.07	7	23
17	-9.12	207.45	4	22
18	-8.16	1.05×10^3	7	28
19	-8.19	993.62	6	25
20	-8.88	311.61	6	31

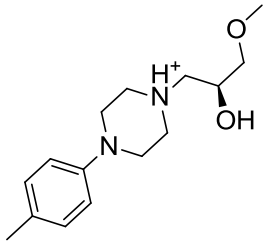
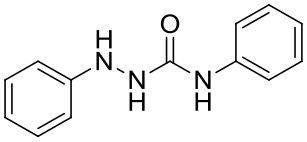
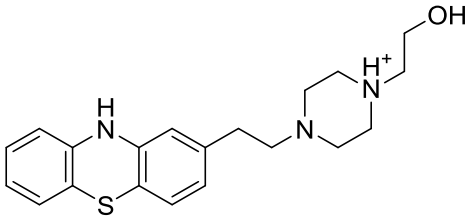
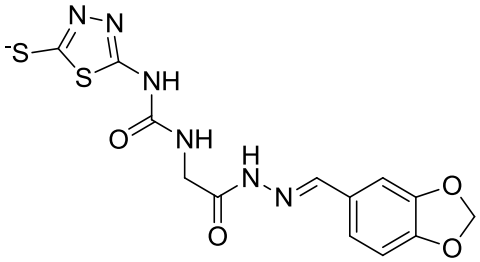
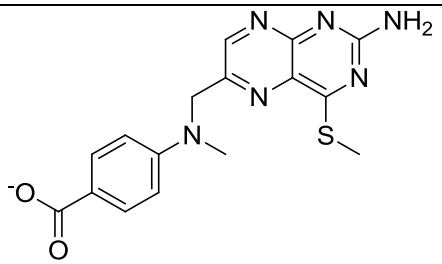
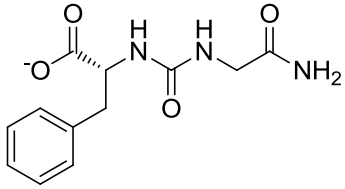
Table 6.7 Top 5 hits identified for site_2 using the *mature glycoprotein receptor* (3N42), showing the molecular weights, calculated logP (clogP), predicted binding energies, inhibitory constants (K_i) and the interaction residues.

	Structure	Molecular weight	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues#
1		376.50	2.37±0.60	-9.98	48.38	E1 Lys52 E1 Thr53 E1 Ile55 E2 Arg36 E2 Glu168
2		354.43	4.12±0.64	-9.71	75.78	E1 Lys52 E1 Ile55 E2 Arg36 E2 Glu168 E2 Tyr237
3		376.50	3.90±0.72	-9.36	138.07	E1 Lys52 E1 Ile55 E2 Tyr237
4		390.50	3.03±0.75	-9.26	163.4	E1 Lys52 E1 Thr53 E1 Ile55 E2 Arg36 E2 Tyr237
5		350.39	2.83±0.90	-9.17	190.7	E1 Lys52 E1 Thr53 E1 Ile55 E2 Arg36 E2 Tyr237

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada). #Numbers of E2 residues in the mature form are different than the corresponding residues in the immature form.

Table 6.8 Fred output virtual screening top 20 hits and their target residues, for the *immature glycoprotein receptor* (3N40) site_4.

Entry	Structure	H-bonding	Other interactions
1		E1 Val 229 x 2 E2 Leu 305	Sigma–π (E2 His 93, x 2) π–sigma (E2 Leu 80)
2		E2 His 82, Leu 305	π– π (E2 His 93)
3			π–sigma (E2 Leu 80)
4		E2 Gln 300, E1 Fusion Loop Trp 98	
5		E1 Fusion Loop Phe 87, Val 229 E2 His 82	π–sigma (E1 fusion loop Met88)
6		E1 Leu 305	
7		E1 Fusion Loop Phe 87 E2 His 82, His 93 x 2	π–sigma (E2 Leu 80)

8		E1 Val 229 x 2 E2 Leu 305 x 2	π–sigma (E2 Leu 80)
9		E1 Val 229, Gly 227, Fusion Loop Phe 87 E2 His 82	
10		E2 His 93	Cation– π (E2 His93) π–sigma (E1 fusion loop Met88) π–π (E1 fusion loop Trp 89) π–sigma (E2 Leu 80)
11		E1 Fusion Loop Phe 87 E2 His 82 x 2	
12		E2 His 82, Asn 136	π–sigma (E2 Leu 80)
13		E2 Leu 305, His 82 E1 Fusion Loop Phe 87 x 2	

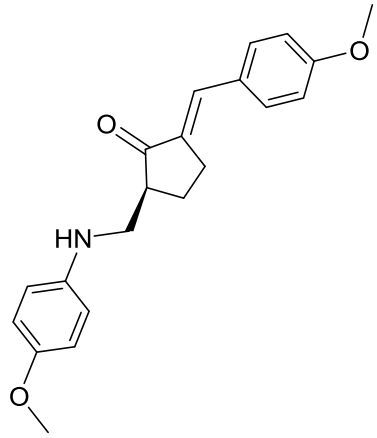
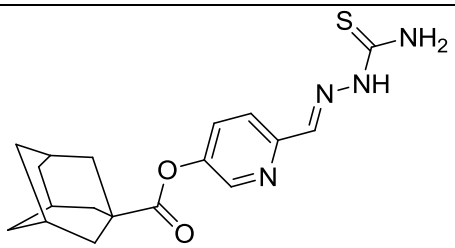
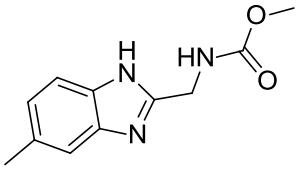
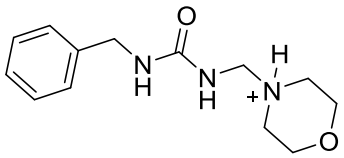
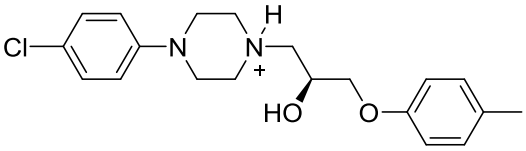
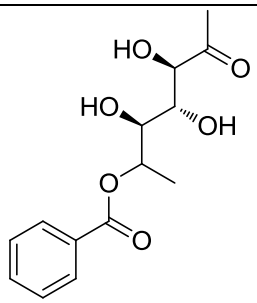
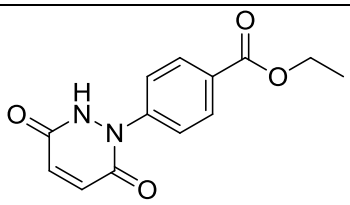
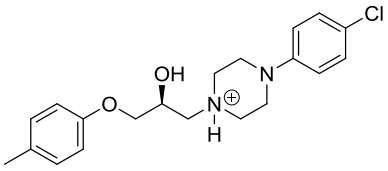
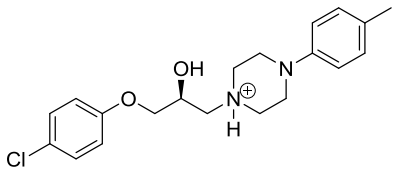
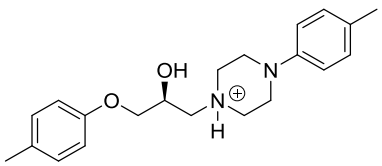
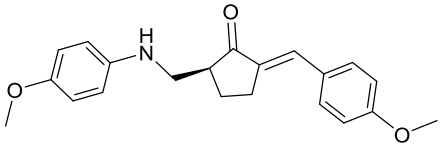
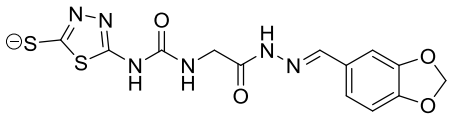
14		Shape Fitting	
15		E1 Thr 228, Fusion Loop Trp 89	
16		E1 Fusion Loop Phe 87 x 2, Gly 227 E2 His 82 x 2	
17		E1 Gly 227, Fusion Loo Phe 87 x 2 E2 His 82, His 93	
18		E2 Leu 305	
19		E1 Fusion Loop Trp 89, Fusion Loop Phe 87, Val 229 E2 His 93	π–sigma (E2 Leu 80)
20		E1 Fusion Loop Trp 89, Val 229	π–sigma (E2 Leu 80) π–sigma (E1 fusion loop Met 88)

Table 6.9 Autodock output for the *immature glycoprotein receptor* (3N40) site_4 Fred hit list, Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0, 3.0 Å.

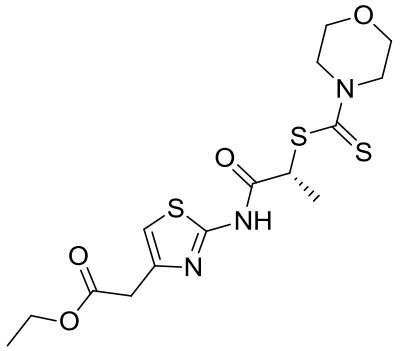
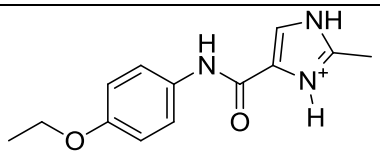
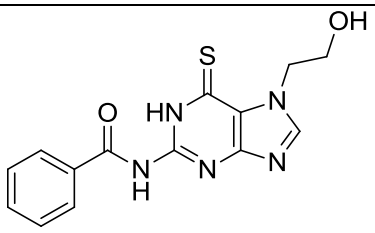
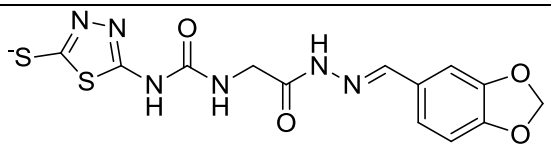
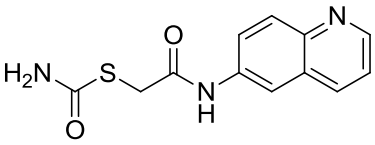
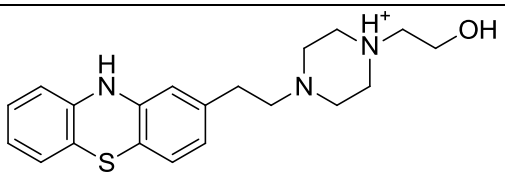
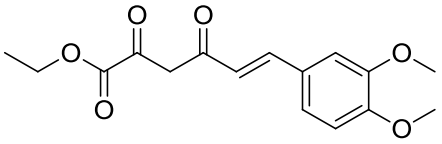
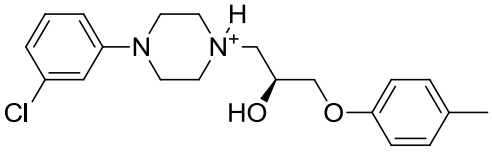
Entry	Binding Energy *Kcal/mol*	Predicted K_i (nM)	Number of clusters	Lowest energy cluster poses
1	-9.47	114.55	8	29
2	-10.07	41.32	3	44
3	-11.23	5.91	4	45
4	-9.78	67.42	16	6
5	-9.79	66.16	3	23
6	-11.2	6.19	4	47
7	-10.03	44.5	4	42
8	-8.64	465.62	2	49
9	-8.39	704.34	3	30
10	-9.64	85.29	6	31
11	-10.45	21.98	5	42
12	-9.49	110.32	6	35
13	-7.31	4.39×10^3	10	23
14	-10.69	14.49	5	33
15	-10.15	36.19	5	26
16	-7.78	2.0×10^3	6	32
17	-7.61	2.62×10^3	8	15
18	-11.3	5.18	6	41
19	-7.42	3.64×10^3	10	17
20	-9.2	181.14	1	50

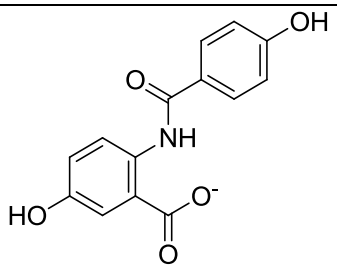
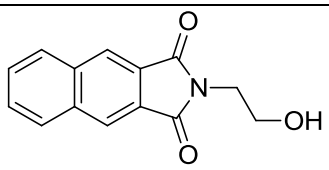
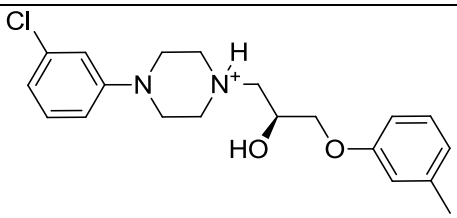
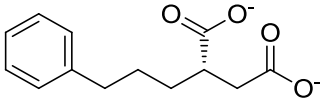
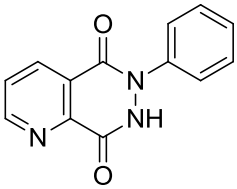
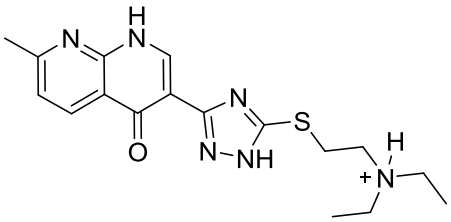
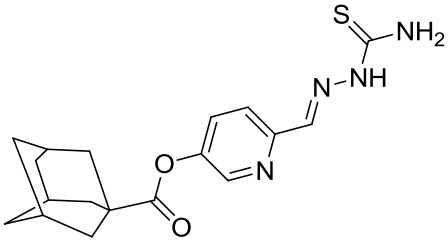
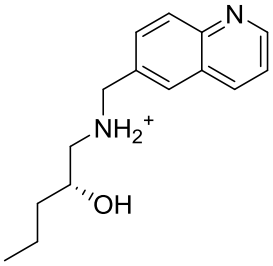
Table 6.10 Top 5 hits identified for site 4 using the *immature glycoprotein receptor* (3N40), showing the molecular weights, calculated logP (clogP), predicted binding energies, inhibitory constants (K_i) and the interaction residues.

	Structure	Molecular weight	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues
1		361.89	4.76±0.54	-11.30	5.18	E1 Val229 E2 His82 E2 His93 E2 Leu80 E2 Leu305
2		361.89	4.71±0.48	-11.23	5.91	E1 Val229 E2 His82 E2 His93 E2 Leu80 E2 Leu305
3		341.47	4.49±0.47	-11.20	6.19	E1 Val229 E2 His82 E2 His93 E2 Leu80 E2 Leu305
4		337.41	3.42±0.34	-10.69	14.49	E2 His82 E2 His93 E2 Leu80
5		379.39	1.83±0.90	-10.45	21.98	E1 Phe87 E2 His82 E2 His93 E2 Ser91 E2 Leu80 E2 Leu305

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada).

Table 6.11 Fred output virtual screening top 20 hits and their target residues, for the *mature glycoprotein receptor* (3N42) site_4

Entry	Structure	H-bonding	Other interactions
1		E2 His 18, His 29	
2		E2 Leu 16, Leu 241 E1 Fusion Loop Trp 89	π–sigma (E2 Leu 16)
3		E2 Arg 13, Thr 175, Leu 16 x 2 E1 Fusion Loop Trp 89	π–sigma (E2 Leu16)
4		E1 Fusion Loop Phe 87 x 2 E2 His 18, His 29	π–cation (E2 His 29)
5		E1 Fusion Loop Trp 89 E2 Arg 13, Leu 16, His 18	π–sigma (E2 Thr 175)
6		E2 His 29	Cation–π (E2 His 29)
7		E2 Arg 13, His 18, His 29 E1 Val 229	
8		E2 His 18, Leu 241	π–cation (E2 His 18) π–π (E2 His 29)

9		E2 Arg 13, Gln 236, Leu 16, E2 His 29 E1 Fusion Loop Trp 89	
10		E2 Arg 13 x 2, Gln236, Thr 175	
11		E2 His 18, Leu 241	π–cation (E2 His 18) π–π (E2 His 29) π–sigma (E2 Leu 16)
12		E2 Leu 16 E2 Arg 13	π–sigma (E1 fusion loop Met 88)
13		E2 Arg 13, Gln 236	π– sigma (E1 fusion loop Met 88)
14		E1 Val 229 x 2 E2 His 29, Leu 241	π–sigma (E2 Leu 16) π–sigma (E1 fusion loop Met 88) Cation–π (E2 His29)
15		E2 His 29, Thr 228	π–cation (E2 His 29)
16		E2 Arg 13, Leu 241, Leu 16	π–sigma (E2 Thr 175)

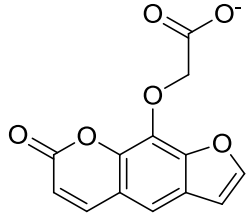
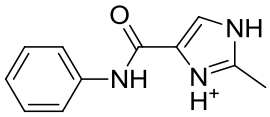
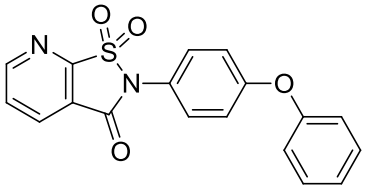
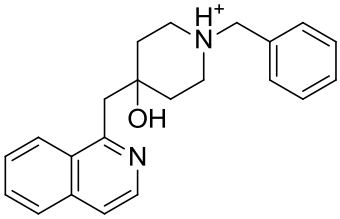
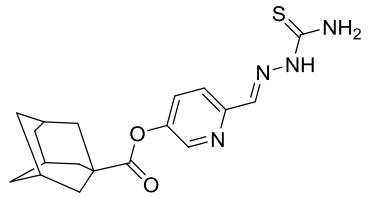
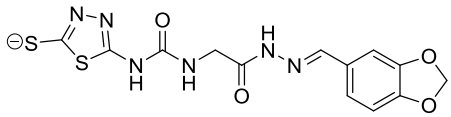
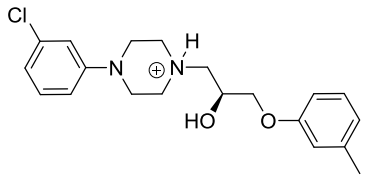
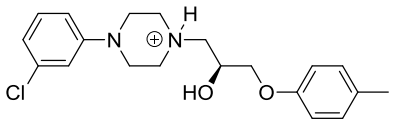
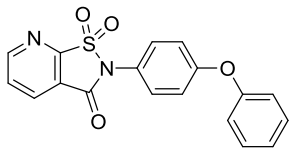
17		E2 Arg 13, Gln 236	π–sigma (E2 Leu 16)
18		E1 Fusion Loop Trp 89 E2 Leu 241, Leu 16	π–sigma (E2 Leu 16)
19		E2 Arg 13	π–sigma (E2 Thr 175) π–cation (E2 His 73) π–cation (E2 His 29)
20		E2 Leu 16	π–cation (E2 His 18) π–cation (E2 His 29)

Table 6.12 Autodock output for the *mature glycoprotein receptor* (3N42) site_4 Fred hit list, Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0, 3.0 Å.

Entry	Binding Energy Kcal/mol	Predicted K_i (nM)	Number of clusters	Lowest energy cluster poses
1	-9.62	88.92	13	17
2	-8.42	678.5	2	45
3	-9.58	95.8	3	47
4	-10.25	30.49	5	36
5	-8.96	270.32	4	37
6	-8.76	380.09	10	21
7	-8.78	368.52	4	31
8	-9.98	48.35	4	43
9	-9.63	87.87	1	50
10	-8.54	551.47	1	50
11	-10.0	46.61	6	39
12	-7.75	2.09×10^3	2	49
13	-8.78	366.14	1	50
14	-9.48	112.05	7	23
15	-10.91	10.03	3	43
16	-7.87	1.72×10^3	4	38
17	-7.32	4.29×10^3	1	50
18	-7.51	3.13×10^3	2	2
19	-9.88	57.61	2	42
20	-8.89	302.63	5	17

Table 6.13 Top 5 hits identified for site 4 using the *mature glycoprotein receptor* (3N42), showing the molecular weights, calculated logP (clogP), predicted binding energies, inhibitory constants (K_i) and the interaction residues.

	Structure	Molecular weight	clogP*	Binding energy *Kcal/mol*	Predicted K_i (nM)	Interaction residues#
1		358.46	3.64±0.39	-10.91	10.03	E1 Thr228 E1 Gly229 E2 His18 E2 His29
2		379.39	1.83±0.90	-10.25	30.49	E1 Phe87 E2 His18 E2 His29 E2 Ser27 E2 Leu16 E2 leu241
3		361.89	4.98±0.48	-10.00	46.61	E1 Val229 E2 His18 E2 His29
4		361.89	4.98±0.48	-9.98	48.35	E1 Val229 E2 His18 E2 His29
5		352.36	2.47±1.22	-9.88	57.61	E1 Trp89 E2 His18 E2 His29 E2 Leu16

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada). #Numbers of E2 residues in the mature form are different than the corresponding residues in the immature form.

6.3 Analysis of the docking calculations

The essential role of the CHIKV envelope protein in the fusion process, its location on the surface of the mature virus (spikes) and the availability of the crystal structures make it a suitable target for structure-based drug design. The CHIKV glycoprotein exists in two forms, the immature form and the mature form. The immature form represents the early stages of the envelope protein after the replication cycle, translated in the endoplasmic reticulum and processed in the Golgi for maturation, moved to the plasma membrane, where it is cleaved by furin-like protease activity in the host infected cell into E2 and E3.[42] The furin cleavage occurs at the furin loop, which represents the junction between E2-E3. The difference between the two crystal structures is the removal of the furin susceptible peptide motif that resulted in slight changes in the volumes of the predicted binding sites. We searched for possible binding sites within both the immature and the mature crystal structures. Six common sites were detected in both structures (Table 1). Among the detected sites, site 2 (the light green colour in Figure 6.2) and site 4 (the blue colour in Figure 6.2) were interesting. Site 2 represents a surface cavity that lies between the E1 domain II and E2 β–ribbon that connects E2 domain A to E2 domain C; it also extends downwards as a channel between E1 domain II and E2 domain A. E2 domains A and B move relative to each other in the pre- and post-fusion structures. Therefore, small molecules that bind to this site may stabilize the E1-E2 heterodimer and prevent their dissociation during the fusion process. Another hypothesis is that it may also stabilize the orientation of E2 domain A with respect to domain B in a way that inhibits the exposure of the fusion peptide in conditions of low pH in the endosome, preventing the fusion process. Moreover, being a groove in this area looking like the enzyme mouth (Figure 6.4), bound small molecules in this site might act as indirect allosteric inhibitors for the furin susceptible peptide motif, and therefore, might impair the cleavage step by the furin proteases. The indirect allosteric inhibition mechanism might be through the inhibition of the interaction between the CHIKV envelope protein, and hence the furin susceptible motif (furin loop), and the acting protease (therefore, the Life Chemicals protein-protein inhibitors library was used here), or through trapping the glycoprotein conformation in one inactive form (relative to the furin cleavage step), which does not interact with the acting proteases. The site 2 volume in the immature form is 9.5% bigger than that in the mature form.

Therefore, both structures were included in the virtual screening study in an attempt to find positive hits for this site.

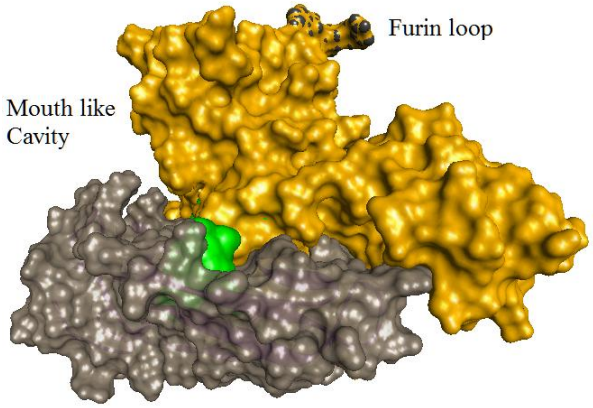


Figure 6.4 Surface representation of the location of site 2 (green), parts of E2 and E3 are shown in orange where the furin loop takes a greyish orange at the top, E1 domain II surface is shown in grey. Site 2 is located in a mouth like cavity that might interact during the furin cleavage. Generated from the PDB file: 3N40.[95]

Site 2 makes close contact with residues from E1 and E2. The E1 residues are: Glu50-Val60, Val229-Pro237; the E2 residues are: Ala97-Arg102 (corresponds to Ala33-Arg38 in the mature form), Gln300-Arg308 (Gln236-Arg244 in the mature form). Hydrogen bonding within these residues involve E1 Lys52, Thr53, Ile55, Val231, His230, and E2 Tyr301 (Tyr237 in the mature form), Glu232 (Glu168 in the mature form), and Arg100 (Arg36 in the mature form). Val54, Lys52, Arg100, Ile167 are able to form other types of strong noncovalent molecular interactions. Generally, valine, alanine and proline amino acids within this pocket are also able to participate in the hydrophobic interactions.

Surprisingly, we could not find a common hit ligand that fits in site 2 in both the immature and the mature forms. However, inspection of the top 5 docked poses in each site reveals that they have the common sequence: heterocycle-S-CH_2-CO-N, the amidic nitrogen in this sequence might be NH, and also can be a part of another ring system. Figure 6.5 shows the 2D representations of the top docked poses in site 2 for both the immature and the mature forms of the envelope glycoproteins. The presence of an electron rich system results in strong noncovalent molecular interactions such as the π-

cation interaction with E2 Arg100 (Arg36 in the mature form) and with E1 Lys52. The heterocyclic ring adjacent to the sulphur in most of the top ranked poses can accept H-bonding with E1 Lys52, Ile55, and Thr53. Being able to bind to residues in both E1 and E2, the ligands identified for this site are most likely to confirm our hypothesis and stabilize the E1-E2 heterodimer and prevent the dissociation.

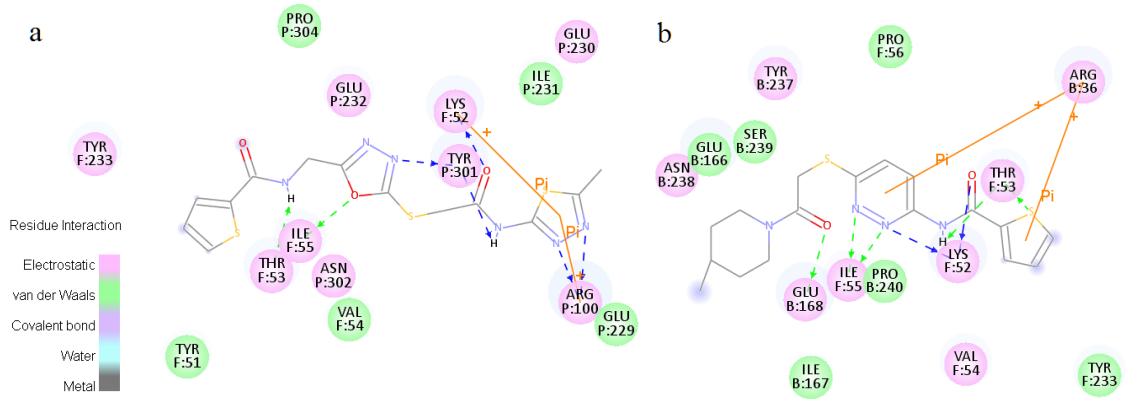


Figure 6.5 2D representation of the docked poses within the binding site 2. (a) Top ranked pose (number 1 in Table 6.4) within the *immature glycoprotein* complex site 2. (b) Top ranked pose (number 1 in Table 6.7) within the *mature glycoprotein* binding site 2. H-bonds are shown in green and blue dashes, while π interactions are shown as orange lines.

Site 4 (blue colour in Figure 6.2) can be described as a narrow channel extending just behind the fusion loop and surrounded by both E2 domains A and B. Comparison between the two sites in the immature and the mature forms indicates that the site volume is 31.3% bigger in the mature form (Figure 6.2), and looks like an umbrella above this site in the immature form. Small molecules binding to this narrow channel will have significant effects; this might not only freeze the relative movement of E2 domains A and B, but might also freeze the fusion loop through stabilizing interactions, and consequently, prevent the exposure of the fusion loop. The fusion loop is stabilized by the histidine residues of E2, which act as the pH sensors for the activation of the fusion protein at lower pH.[95] This site cavity lies in contact with several histidine residues of E2. Therefore, blocking this site may also impair the pH sensor activation mechanism. All the identified hits were found to bind to both E1 and E2 residues, involving the histidine residues of E2, moreover, two hits were found to

bind to the fusion loop amino acids, confirming the ability of freezing the fusigenic activity of the envelope proteins.

Site 4 forms close contact with the E1 fusion loop residues Pro86-Gly91, E1 Gly227-His230. The fusion loop Gly91 and His230 (outside of the fusion sequence) were found to be critical for the fusion.[96] This emphasizes the importance of our hypothesis that binding to this site will impair the fusion process. From E2, residues Arg77-His82 (Arg13-His18 in the mature form), Ser91-Val96 (Ser27-Val32 in the mature form) and residues Leu305-Ala310 (Leu241-Ala246 in the mature form) form close contacts with the binding site.

Interestingly, a common hit ligand was found in both the blue sites (ranked 5 in Table 6.10, and ranked 2 in Table 6.13). It shows the same interactions within the two binding pockets and more importantly, it forms two H-bonds (5.4, 6.3 Å) with E1 fusion loop amino acid Phe87 backbone carbonyl, three amino acids away from Gly91, the critical[96] residue for the fusion process. It also binds to E2 His82, His93 (His18, His29 in the mature form) *via* H-bonding and π-cation interactions (Figure 6.6). Moreover, the predicated binding affinity and inhibitory constant (K_i, in the nano molar range), along with the acceptable clogP value of 1.8 (Tables 6.10, 6.13), make it a very attractive candidate for developing anti-chikungunya drug targeting the envelope proteins.

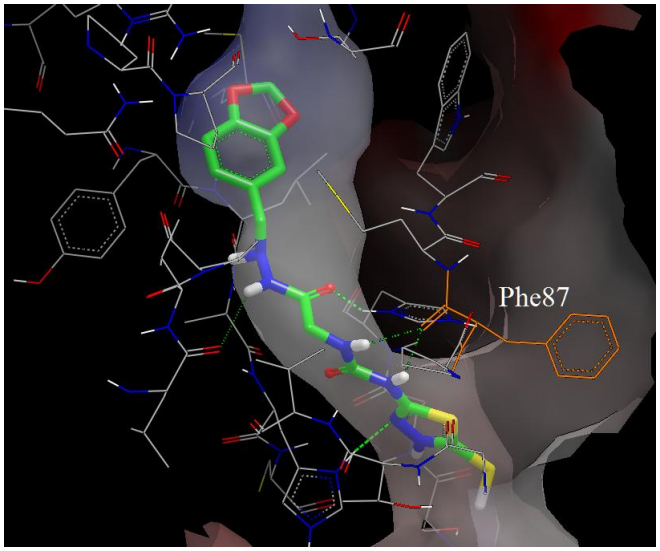


Figure 6.6 3D representation of the predicted docking pose of compound 2 (Table 6.13) within site 4 binding pocket (transparent surface) of the *mature* chikungunya envelope protein showing the H-bonding (green dashes) with the fusion loop amino acid Phe87 (shown in orange colour).

A further interesting observation is the presence of same chiral skeleton; the (*S*)-1-(2-hydroxy-3-phenoxypropyl)-4-phenylpiperazin-1-ium in a series of compounds (1, 2 and 3 in Table 6.10) and (3, 4 in Table 6.13). The compounds only differ in the substituents on the terminal phenyl rings. The chirality of this series indicates the selectivity of the compounds and reflects the importance of the stereochemistry in designing inhibitors for this site. The enantiomers of these compounds (within the library) did not pass the first FRED virtual screening. Within the immature narrow binding pocket (site 4), these series were able to form H-bonds with the E1 Val229, E2 His82, E2 Leu305. However, in case of the 31.3% bigger pocket of the mature site 4, this series was able to achieve the H-bonding with E1 Val229, E2 His29 whereas it failed to form H-bonds with the E2 Leu241, but was still able to achieve the π stacking interaction with E2 His18 and His29 (Figure 6.7). The importance of this stereo-selectivity in inhibiting the envelope protein, was also noticed recently in the inhibitors of the dengue virus envelope proteins mediated fusion, where compounds with certain stereochemistry of the OH group (the (*S*) enantiomers) were shown to have stereospecific effects on the activities.[237] The importance of the (*S*) configuration of the compounds (Figure 6.7) can be related to the ability of the OH groups of the compounds

to achieve H-bonds with the E2 histidine residues (His82 of the immature form and His29 of the mature form), whereas these H-bonds might not be possible with the other enantiomers. Superimposition of the two compounds (number 1 in Table 6.10 and number 3 in Table 6.13) within the binding pockets (site 4 in both enzyme forms) revealed the positioning of the docked poses in a similar way with slight changes in the orientation of the hydroxyl groups and the central piperazinium moiety (Figure 6.8). It is also clear that the chlorine atoms seem not essential for interactions. This superimposition not only indicates the reliability of the interactions of this class of compounds with the residues within this site in the two forms of the enzyme, but also confirms our hypothesis that this series might be developed as selective CHIKV envelope protein inhibitors.

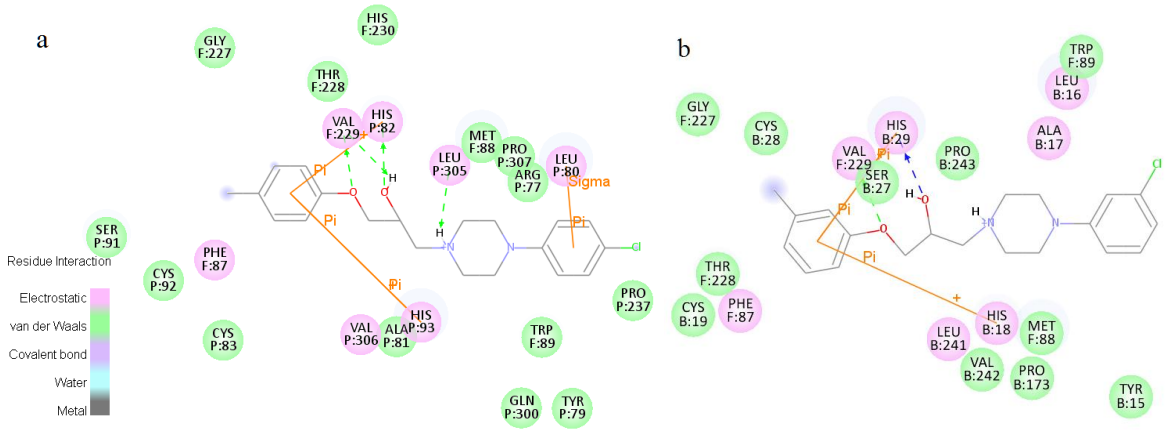


Figure 6.7 2D representation of the docked poses within the binding site 4. (a) Top ranked pose (number 1 in Table 6.10) within the *immature glycoprotein* complex site 4. (b) Top ranked pose (number 3 in Table 6.13) within the *mature glycoprotein* binding site 4. H-bonds are shown in green and blue dashes, while π interactions are shown as orange lines.

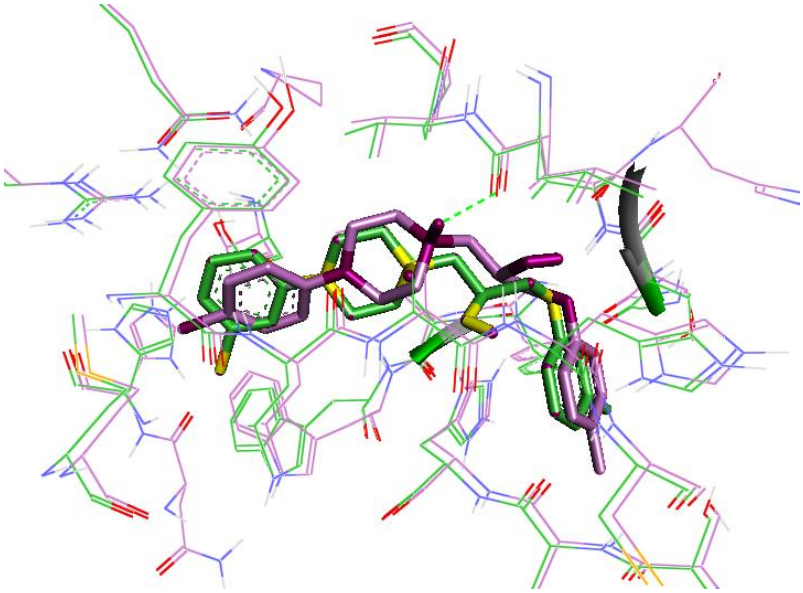


Figure 6.8 Superimposition of compound 1 (in Table 6.10) within the *immature glycoprotein* site 4 (compound and residues are shown in violet colour) and compound 3 (in Table 6.13) within the *mature glycoprotein* binding site 4 (compound and pocket residues are shown in green colour). Slight differences can be observed for the orientation of the hydroxyl groups and the central piperazinium ring.

Although ranked 5 in Table 6.13, this ligand shows extraordinary H-bonding (2.41 Å) with the fusion peptide amino acid Trp89 backbone (Figure 6.9), just one amino acid away from Gly91, the critical[96] amino acid for the fusion process, The pose is also stabilized inside the pocket by the interactions with the E2 His18, His29 and Leu16 (Table 6.13). This also emphasizes the possibility of inhibiting the fusion process through designing ligands for this pocket.

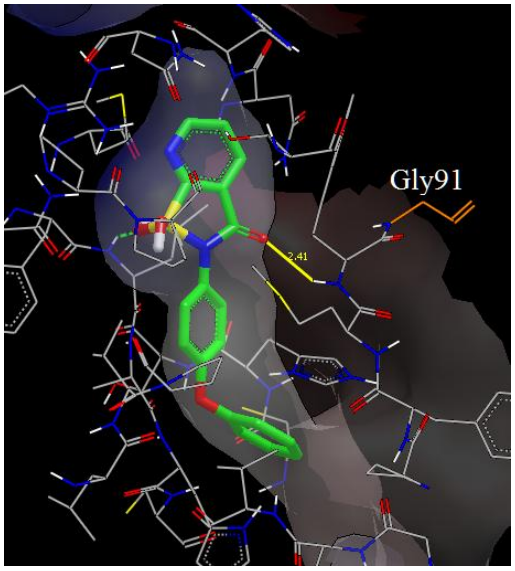


Figure 6.9 Pose 5 (Table 6.13) within the *mature glycoprotein* binding site 4, showing the 2.4 Å H-bond interaction (yellow line) with Trp89. Gly91 is shown in orange.

6.4. Concluding remarks

Thus far, the CHIKV envelope protein has not been investigated as a possible target for the drug design against the virus. Therefore, we have investigated for the first time the possible binding target sites within the immature and the mature forms of the CHIKV envelope proteins. We managed to identify two sites that look critical to the protein functions; mainly the fusion process, based on the functionality and the location of the sites. A virtual screening search was performed on the two sites in both forms of the enzymes to increase the chances of finding reliable positive hits. Five hits for each site in both forms of the CHIKV envelope proteins were identified revealing some important features for further developing antagonists for these proteins. To test our hypothesis, the identified hits need to be evaluated against the CHIKV, which is currently under investigation. Our study represents a good template for designing selective inhibitors for the chikungunya virus envelope proteins *via in silico* and *in vitro* optimization process. This developed computer model can be further used for screening larger database libraries such as the FDA approved drugs libraries that would facilitate the identification of effective therapies against the CHIKV. Our hypothesis might also be a useful tool for inhibiting other *alphaviruses* such as Sindbis virus and Semliki Forest virus as well as other fusion mediated viruses.

CHAPTER 7: New Leads for African Trypanosomiasis

7.1. Introduction

Trypanosoma brucei gambiense and T. b. rhodesiense are the causative agents of the Human African Trypanosomiasis (HAT), also known as sleeping sickness. Anothere neglected disease as the Chikungunya virus, the major topic of this thesis. *Trypanosoma brucei rhodesiense* is found in Eastern and Southern Africa, whereas *Trypanosoma brucei gambiense* occurs in Western and Central Africa and is responsible for over 90% of all reported cases of infection.[238] This disease threatens about 60 million people living in sub-Saharan Africa and causes an estimated 25,000 deaths per year.[239] It has a major impact on the affected nations causing suffering and poverty and if left untreated, the disease is usually fatal.[240] The lack of full-scale screening programmes and poor diagnostic tools has led to under-reporting of cases, which is likely to be at least threefold higher than the measured value.[241]

Both subspecies are transmitted by the bite of the infected tsetse fly. *T. b. gambiense* HAT is primarily a chronic human disease whereas *T. b. rhodesiense* HAT is primarily zoonotic with a huge animal reservoir, causing the acute form of sleeping sickness. After the bite, the parasite starts to multiply in the blood (phase I, febrile illness with flu-like symptoms, which can last for several weeks or months). During this phase, the parasite lives within the bloodstream and subsequently migrates to other areas of the human body, such as the lymph nodes, spleen, and spinal fluid, causing symptoms similar to those caused by malaria (rash, fever, shaking chills, body aches, and general fatigue). If phase I is left untreated, the parasite then crosses the blood brain barrier attacking the central nervous system (phase II) and neurological symptoms start to appear, causing progressive mental deterioration, sleep disturbances, long lasting coma (disruption of the sleep/wake cycle and psychological effects) and finally death if not treated.[242]

Unfortunately, vaccines are not available for this disease. The unavailability of vaccines is mainly due to the high degree of antigenic variation expressed by the glycoprotein forming their surface coat.[243] Therefore, the main defence against the

parasite would be chemotherapeutics. The treatment options are limited with only few drugs available. The four registered drugs to treat HAT, are suramin **152** and pentamidine **153** for early stage infections, whereas melarsoprol **154** (contains arsenic) and eflornithine **155** are for the late-stage treatment (Figure 7.1).[244,245] All these drugs were developed approximately 30 years ago and suffer high toxicity, the emergence of resistance and lack of efficacy. Melarsoprol **154** is the most toxic and causes reactive encephalopathy in 5−10% of treated patients, with 1−5% mortality.[246]

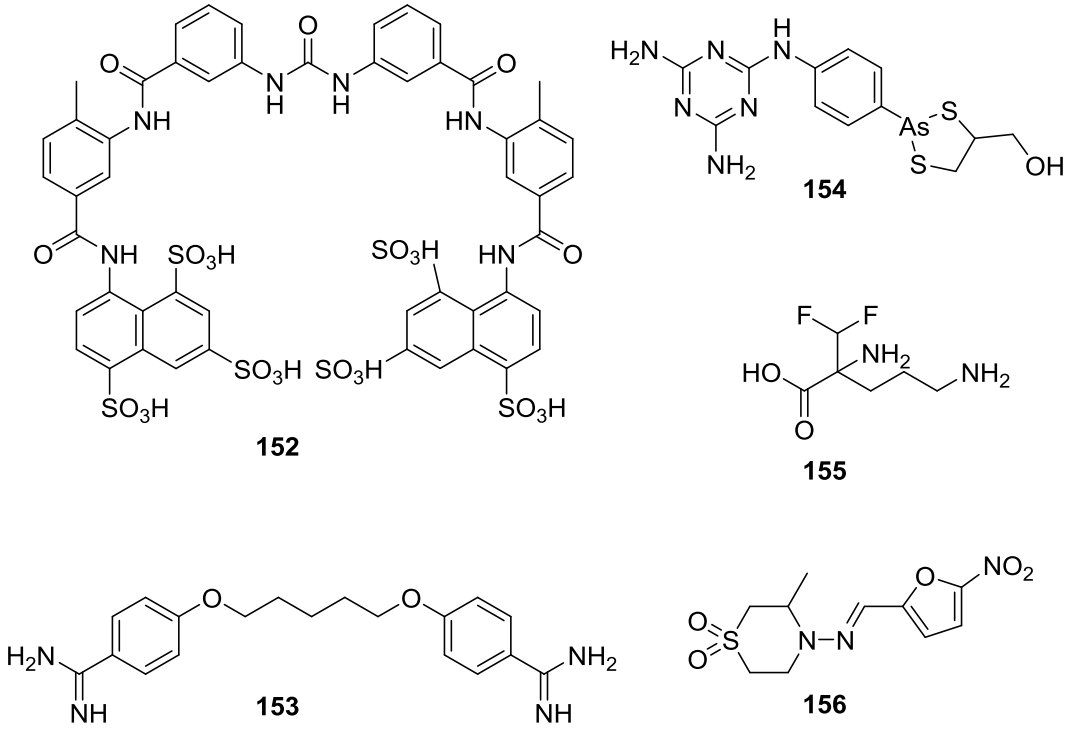


Figure 7.1 Older generation of African sleeping sickness treatment drugs.

Recently, there has been some progress in the treatment of HAT,[247] with the nifurtimox **156**−eflornithine **155** (Figure 7.1) combination (NECT) chemotherapy developed to reduce resistance frequency.[248,249] However, these options are still under investigation due to inactivity against some trypanosome species or phases of the infection, as well as toxicity, the parenteral mode of administration, and the emergence of resistance.[250,251]

Other trypanocidal agents that are active against the first stage of HAT and are orally bioavailable are pafuramidine **157**[252] and fexinidazole **158**[253] (Figure 7.2).

Although it was used in phase III clinical trials in 2007, pafuramidine **157** use was terminated in 2008 because of hepatic toxicity. More recently, the novel boron-containing molecule **159** (Figure 7.2) emerged as an orally active drug candidate that completed preclinical studies in 2011.[254] It showed promising activity against the blood and the brain parasite. The mechanism by which compound **159** acts as a trypanocidal agent is still unknown. It entered phase I clinical trials in France in 2012 for evaluation against phase 1 and phase 2 HAT. The present treatment options indicate that new therapeutics with acceptable efficacies and safety profiles are urgently needed new therapeutic opportunities.[255]

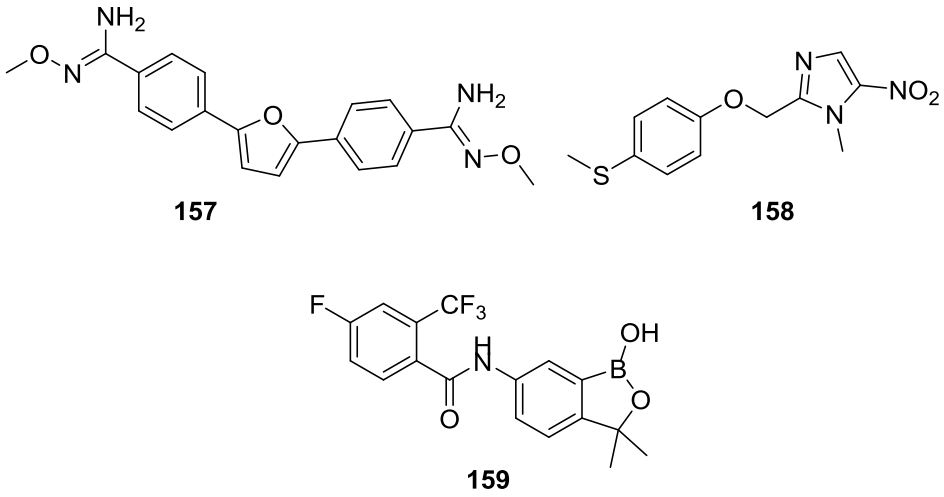
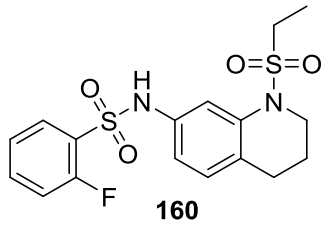


Figure 7.2 New generation of African sleeping sickness treatment promising agents.

High throughput screening (HTS) is one approach that can be used to identify new lead compounds for such neglected diseases. Therefore, in a previous work,[256] the HTS library of 87,926 compounds (WEHI 2003)[256] was tested against the related trypanosome subspecies, *Trypanosoma b. brucei* and against a mammalian cell line HEK293, to determine a selectivity index for each compound. Cluster analysis, considering chemical alerts such as toxicophores, the likelihood of CNS penetration, and drug-like structural features yielded a subset of twelve compounds as promising medicinal chemistry starting points for drug development. A consortium of Australian medicinal chemists was formed among different Australian universities, to optimize the obtained hits, and some of these hits were introduced to our group for optimization.

This chapter discusses the synthesis and anti-trypanocidal activity of new analogues for the bis-sulfonamide hit, WEHI-1203255 (**160**, Figure 7.3), which showed an IC_{50} 1.3 µM, with a selectivity index (SI) of >32. This hit compound **160** has excellent physiochemical properties, good calculated aqueous solubility of 100 µM, an acceptable polar surface area of 84 $Å^2$, and an acceptable CLogP value[231] of 2.5. The analogues were synthesized and tested for their ability to inhibit the growth of *T. brucei* limiting the changes to the two sulfonamide moieties to study the preliminary structure activity analysis. The cytotoxicity profiles of the compounds were evaluated using (HEK293) cell line and selectivity index was estimated for each analogue. The SI of the compounds was determined where possible by directly comparing the IC_{50} values from the *T. b. brucei* and (HEK 293) assay. If this was not possible, an estimated IC_{50} value was calculated by comparing the dose at which the compound was active >50% in the *T. b. brucei* assay and the highest dose at which there was no activity (<50%) in the (HEK 293) assay.



160

Figure 7.3 Lead compound **160** (WEHI 1203255), IC_{50}= 1.3 µM.

7.2. Results and discussion

7.2.1. Synthesis of the bis-sulfonamide analogues

The retro-synthetic analysis of the lead compound **160** (Figure 7.4) revealed that it could be accessed through three facile synthetic steps: the sulfonation of the commercially available 7-nitrotetrahydroquinoline **161**, followed by the reduction of the nitro group to afford the aminotetrahydroquinoline **163**. A second sulfonation of the amino group would result in the final bis-sulfonamide compound **160** (Figure 7.4). This strategy enabled us to access 26 separate derivatives in a short timeframe, reliable yields and at reasonable cost.

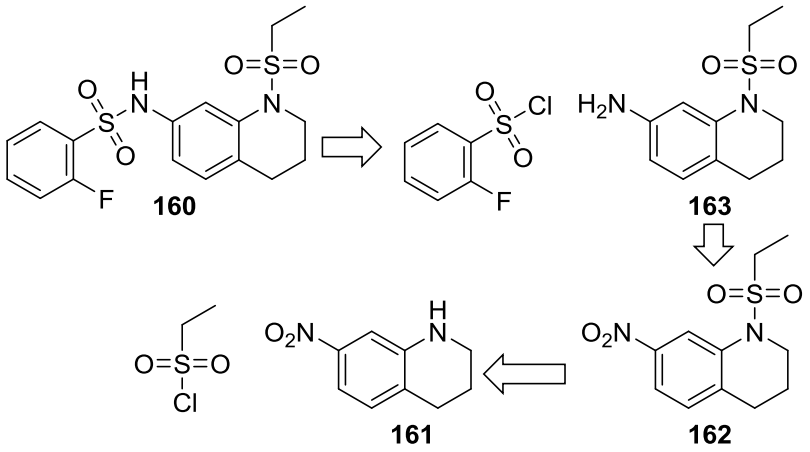
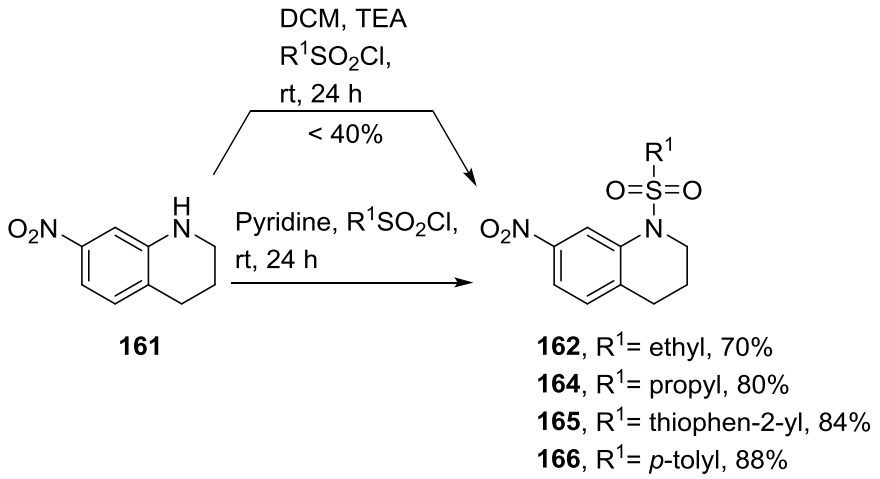


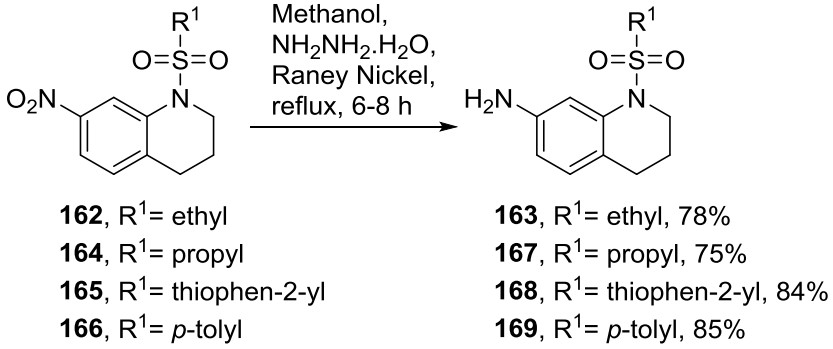
Figure 7.4 The retro-synthesis analysis for the access of the lead compound **160**.

The first step in the synthetic pathway was the sulfonation of the secondary amine of commercially available 7-nitrotetrahydroquinoline **161** using the appropriate sulfonyl chloride. The reaction (Scheme 7.1) was initially attempted using CH_2CH_2 as a solvent in the presence of TEA as a base and stirring the reaction mixture at room temperature for 24 h, however, the yields of the isolated compounds were less than 40%, with the remaining being un-reacted starting 7-nitrotetrahydroquinoline **161**. Pyridine behaves as a better solvent and a stronger base for the tetrahydroquinoline sulfonation,[257,258] and was therefore used throughout in this reaction. Therefore, the 7-nitrotetrahydroquinoline **161** was dissolved in anhydrous pyridine, and the mixture was then stirred at 0 ºC. The appropriate sulfonyl chloride (liquid) was then added dropwise and the mixture was stirred for 24 h (Scheme 7.1). The reaction mixture was then quenched with cold water, after which the resultant oil or solid was extracted using ethyl acetate. Column chromatography afforded the desired sulfonamides **162**, **164-166** in 70-88% yield. Analysis of the ^{1}H NMR spectrum of 1-(ethylsulfonyl)-7-nitro-1,2,3,4-tetrahydroquinoline **162** showed quartet peak integrating for two protons and resonating at 3.22 ppm, and another triplet peak, integrating three protons, resonating at 2.95 ppm. These peaks were assigned to the ethyl CH_2 and CH_3 groups, respectively. Analysis of the EI-MS spectrum of **162** showed a peak at 270 *m/z* corresponding to the molecular ion (M+).



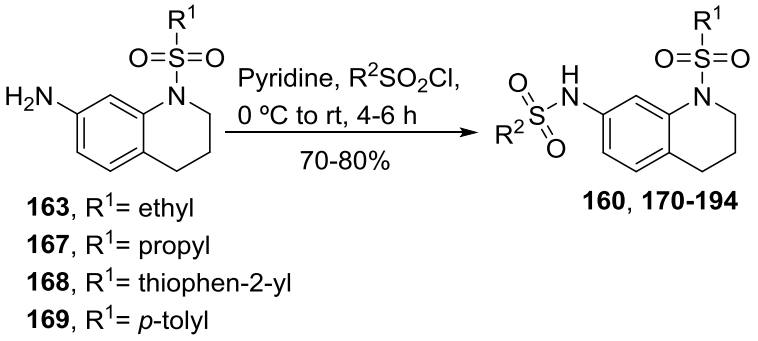
Scheme 7.1 Sulfonation of the tetrahydroquinoline **161** to afford intermediates **162-165**.

The nitro group of the sulfonamides **162**, **164-166** was then reduced, initially using the reduction methodology developed in our lab, where the sulfonamides **162**, **164-166** were dissolved in a mixture of acetic acid, ethanol and water in the presence of iron powder as a catalyst, with sonication at 30 ºC,[209] however, under these conditions, yields of the reduced products were under 50%. Therefore the Raney nickel catalysed reduction was performed by dissolving the nitro derivatives in methanol, followed by the addition of Raney Nickel (as a suspension in water) under a N_2 atmosphere. The mixture was then vigorously stirred (Scheme 7.2) and hydrazine monohydrate, as a hydrogen source, dissolved in methanol was then added dropwise resulting in the production of H_2 gas. After the H_2 gas evolution ceased, the mixture was heated at reflux, filtered hot through celite, and the methanol and excess hydrazine hydrate were removed, affording the aniline derivatives **163**, **167-169** in 80-90% yields, which were easily purified by recrystallization from either methanol or toluene. Analysis of the EI-MS spectrum of 7-amino-1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinoline **163** showed a peak at 240 *m/z* which was assigned for the molecular ion (M+). Analysis of the ^{1}H NMR spectrum of **163** showed a broad singlet peak resonating at 3.62 ppm that was assigned for the NH_2.



Scheme 7.2 Reduction of the nitrotetrahydroquinolines **162**, **164-166**.

The amino groups in the sulfonamides **163**, **167-169** were then reacted with the appropriate sulfonyl chloride (the second sulfonation) in pyridine at room temperature (Scheme 7.3). After aqueous workup, extraction, washing with HCl and evaporation of the solvents, the bis-sulfonamides compounds **160** and **170-194** were isolated and were purified using column chromatography to afford the final analogues in 70-80% yield. Analysis of the EI-MS spectrum of *N*-(1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2-fluorobenzenesulfonamide **160** showed a peak at 398 *m/z* that was assigned as the molecular ion (M+). Analysis of the ^{13}C NMR spectrum of **160** showed a doublet peak at 159.6 ppm (J = 254.9 Hz) that was assigned to the carbon attached to the F atom (compound **160**, Figure 7.4). HPLC analysis of these bis-sulfonamides showed a purity range > 95% for all the synthesized derivatives.



Scheme 7.3 The synthesis of the bis-sulfonamides **160**, **170-194** (see Table 7.1 for R^1 and R^2).

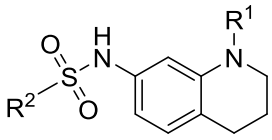
7.2.2. Trypanocidal activity and structure activity relationship (SAR)

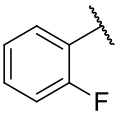
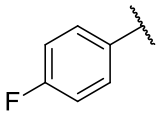
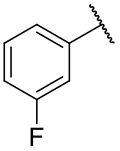
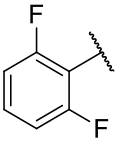
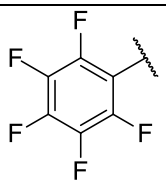
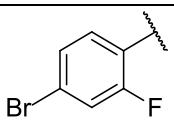
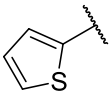
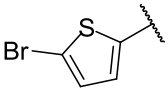
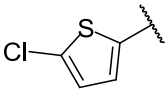
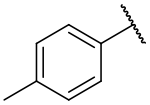
The results for the testing against *Trypanosoma b. brucei*, the calculated ClogP and the selectivity index are listed in Table 7.1. The initial activity was screening at 1 µM and 10 µM and derivatives that showed significant activity were then tested to obtain the IC_{50} values. The first series of derivatives examined the changes in the aromatic sulfonyl moiety where the lead compound **160** has a fluorine atom in the *ortho* position. This lead compound was also resynthesised and tested with IC_{50}/SI= 4.1/>20, confirming the initial activity results from the (WEHI 2003) compounds library. Changing the *ortho*-fluoro substituent to the *para* (compound **170**) and *meta* (compound **171**) positions, did not improve the activity where the IC_{50}/SI profile was 7.8 µM/>10 for both derivatives, indicating decreased trypanocidal activity and increased toxicity. Increasing the number of the fluorine atoms had a negative effect on the activity, where the addition of a second fluoro substituent into the adjacent *ortho* position (compound **172**) resulted in a decreased activity (82% activity at 10 µM) as did the presence of five fluoro substituents (compound **173**, 40% activity at 1 µM). This implied the importance of the mono-fluoro atom only in the *ortho* position, with no advantages with the presence of the extra fluorine atoms.

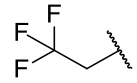
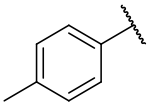
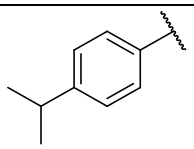
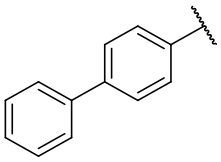
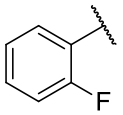
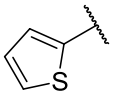
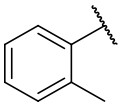
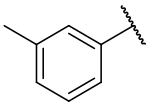
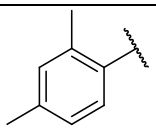
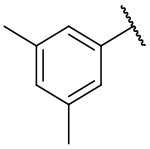
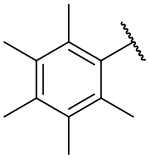
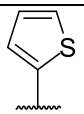
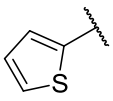
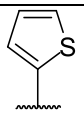
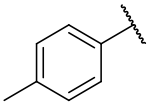
The fluorine atom is the smallest halogen size, therefore, it was interesting to test the presence of other halogen atoms. The addition of a bromo substituent at the *para* position (compound **174**) resulted in increased toxicity and did not improve the activity (IC_{50}/SI= 9.9 µM/>8).

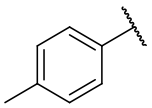
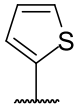
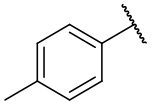
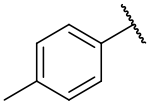
Replacing the benzene moiety with its bioisostere thiophene[259,260] (compound **175**) resulted in a similar activity (IC_{50}/SI= 4.0 µM/>20) compared to the lead **160**. The same activity profile of **160** and the simpler **175** might be attributed to the lipophilic nature thiophene ring in **175** that might have the same effect as the fluorine atom in **160**. The addition of another halogen to this thiophene (Br, compound **176**) showed improved trypanocidal activity (IC_{50}/SI= 2.6 µM/>7), however, the selectivity index value was reduced compared to the unsubstituted thiophene analogue **175**. Interestingly, replacement of the bulky bromo substituent in **176** by a chloro substituent (compound **177**) was not tolerated, as the activity decreased (IC_{50}/SI= 13.1 µM/>6). The activity of **177** was 3 fold less than that of **175** and 5 fold less than that of **176**.

Table 7.1 R^1, R^2, molecular weight, ClogP, IC_{50} and Selectivity index for compounds **160**, **170-19**.



Compound	R^1	R^2	Mwt	ClogP*	IC_{50} (µM)	S.I.
160	—CH_2CH_3		398.47	2.48±0.82	4.1±0.643	20.2
170	—CH_2CH_3		398.47	2.98±0.82	7.8±2.695	10.7
171	—CH_2CH_3		398.47	3.09±0.82	7.8±2.155	10.6
172	—CH_2CH_3		416.46	2.33±0.87	82% @ 10 µM	
173	—CH_2CH_3		470.43	3.37±1.07	40% @ 1 µM	
174	—CH_2CH_3		477.36	3.50±0.88	9.9±3.319	8.4
175	—CH_2CH_3		386.50	2.27±0.79	4.0±0.625	20.5
176	—CH_2CH_3		465.40	3.49±0.87	2.6±2.573	>7
177	—CH_2CH_3		420.95	3.28±0.88	13.1±4.344	6.3
178	—CH_2CH_3		394.50	3.13±0.77	3.4±0.910	24.4
179	—CH_2CH_3		332.43	1.57±0.76	IN	-

180	—CH_2CH_3		386.41	2.82±0.96	IN	-
181	—$CH_2CH_2CH_3$		408.53	3.66±0.77	4.7	>18
182	—$CH_2CH_2CH_3$		436.58	4.54±0.77	3.9	>22
183	—$CH_2CH_2CH_3$		470.60	4.85±0.79	13.0	>6.6
184	—$CH_2CH_2CH_3$		412.49	3.01±0.82	10.0	>8
185	—$CH_2CH_2CH_3$		400.53	2.80±0.79	IN	-
186	—$CH_2CH_2CH_3$		408.53	3.66±0.77	IN	-
187	—$CH_2CH_2CH_3$		408.53	3.66±0.77	IN	-
188	—$CH_2CH_2CH_3$		422.56	4.12±0.77	16.0	>5
189	—$CH_2CH_2CH_3$		422.56	4.12±0.77	IN	-
190	—$CH_2CH_2CH_3$		464.64	5.50±0.78	11.4	>7
191			440.57	2.26±0.83	IN	-
192			448.57	3.12±0.81	3.1±0.893	2.45

193			448.57	3.12±0.81	7.6±2.552	1.5
194			456.57	3.99±0.79	3.9±0.887	1.9

*Calculated using ACDLabs v.12.0 (ACD/Labs, Toronto, Canada).

Replacing the aromatic moiety with aliphatic chains as in compounds **179** and **180** completely abolished the activity which indicates the importance of the aromatic ring in that position for the trypanocidal activity. The π system of the aromatic ring may be involved in the interaction site, whereas with aliphatic side chains, such interactions don't exist. Interestingly, the introduction of a small hydrophobe such as a methyl group on the *para* position on the phenyl ring of the aromatic sulfonyl moiety (compound **178**) improved the activity slightly (IC_{50}/SI= 3.4 μM/>24), mostly due to the increased difference between the effective and toxic concentrations compared to **160**. As the mode of action and the target of the compounds are not yet known, the role of this hydrophobe methyl group cannot be confirmed, however, it might be involved in a hydrophobic interaction within the target site.

Addition of another hydrophobe (a methyl group) on the ethyl side chain of the other sulfonyl group (second changeable moiety) as in compound **181**, was also tolerated (IC_{50}/SI= 4.7 μM/>18) where no significant difference in the activity compared to **178** was observed with a slight narrowing of the selectivity index compared to **178**. In the case of compound **185**, this extra hydrophobe on the ethyl side chain was not tolerated and activity dropped. When the *p*-tolyl group of **181** was replaced by a thiophene ring (compound **185**), the activity was completely abolished due to this extra hydrophobic methyl (the propyl aliphatic group), compared to **175**.

The *para* position of the methyl hydrophobe in **181** was important for activity as can be indicated by the inactivity of compounds **186** (*ortho* position) and **187** (*meta* position). The weak activity of **188** and **190** and inactivity of **189** (no *p*-methyl group) also gives indications about the importance of the *para* position, with no advantages with the presence of the extra hydrophobic methyl groups. Figure 7.5 shows the SAR for this series of bis-sulfonamides.

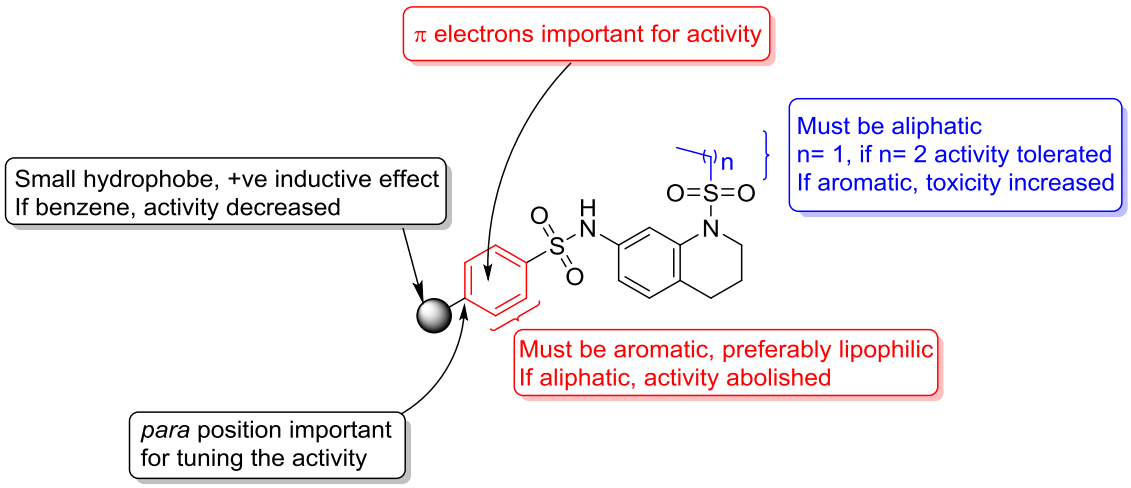


Figure 7.5 Structure activity relationship for the bis-sulfonamides.

Addition of the hydrophobic methyl groups on either the ethyl side chain or on the aromatic sulfonyl moieties seemed to act as a tuner for the activity. Compound **185** was completely inactive whereas compound **184** showed similar activity (IC_{50}/SI= 3.1 µM/>8) compared to the lead **160**, with a narrowing in the selectivity index observed. Increasing the bulkiness of compound **181** to **182** (replacing the small methyl hydrophobe with the more bulky isopropyl group) resulted in an improved activity profile and decreased the toxicity (IC_{50}/SI= 3.9 µM/>22). Increasing the bulkiness by replacing the isopropyl in **182** by a phenyl ring (compound **183**) decreased the activity (IC_{50}/SI= 13 µM/>7) by 3 fold compared to **182**, but did not completely abolish the activity. This difference can be attributed to the difference of the electronic effect; the methyl and isopropyl group have a positive inductive effect on the benzene ring while the phenyl ring at the *para* position has a negative inductive effect. This also confirms that the π system of the aromatic ring (directly attached to the sulphonamide group) might be involved in the activity.

When the aliphatic sulfonyl side chain (sulfonyl group attached to the tetrahydroquinoline N) was replaced with aromatic ring, the activity was either completely abolished as in case of compound **191** or toxicity was increased as in compounds **192-194**. Interestingly, compound **192** carrying the *para* methyl group was the best in this series (IC_{50}/SI= 3.1 µM/>2), confirming the importance of the *para*

position on this aromatic moiety. The aliphatic sulfonyl groups directly attached to the tetrahydroquinoline ring is important for activity rather than an aromatic replacment.

7.3. Concluding remarks

In this chapter, we synthesized 26 bis-sulfonamide analogues of the lead compound **160**, which had good physiochemical properties and was discovered during the high throughput screening (HTS) of the WEHI 2003 library of 87,926 compounds. The analogues were accessed through three facile synthetic steps. Some of the analogues displayed good activity profile ranging from 2 to 4 μM. The preliminary structure activity relationship revealed that the presence of an aliphatic side chain on the tetrahydroquinoline **160** is important for activity and should not be replaced by aromatic groups. The second sulfonation of the amino group of the sulfonated tetrahydroquinoline **166-169** must be with aromatic sulfonyl groups, as when aliphatic groups were used, activity was completely abolished. The bis-sulfonamide **176** with the bromo-thiophene group, displayed IC_{50} value of 2.6 μM with a selectivity index >7. The *para* position of the aromatic ring on the bis-sulfonamides **178**, was found to be critical for tuning the activity, where the activity changed with changing the size of the hydrophobe substituents in that position. Small hydrophobe groups were favoured over the bulky aromatic groups in that *para* position. Therefore, this position should be investigated more with a variety of possible substitutions. The target site and mode of action of the bis-sulfonamides are still yet unknown, and need further investigations.

8.1. Chapter 2: The Search for Anti-CHIKV Lead Compounds

8.1.1. Conclusions

The modifications in the AAPM series (Figure 8.1) were advantageous in decreasing the steric hindrance around the exchangeable chloro substituent, and also in simplification of the synthetic steps required throughout the synthesis. The alkene **63**, with the less hindered cyclopropyl group, represents the most promising alkene to perform more amination reactions under moderate conditions, as can be indicated by the formation of the morpholine analogue **77**.

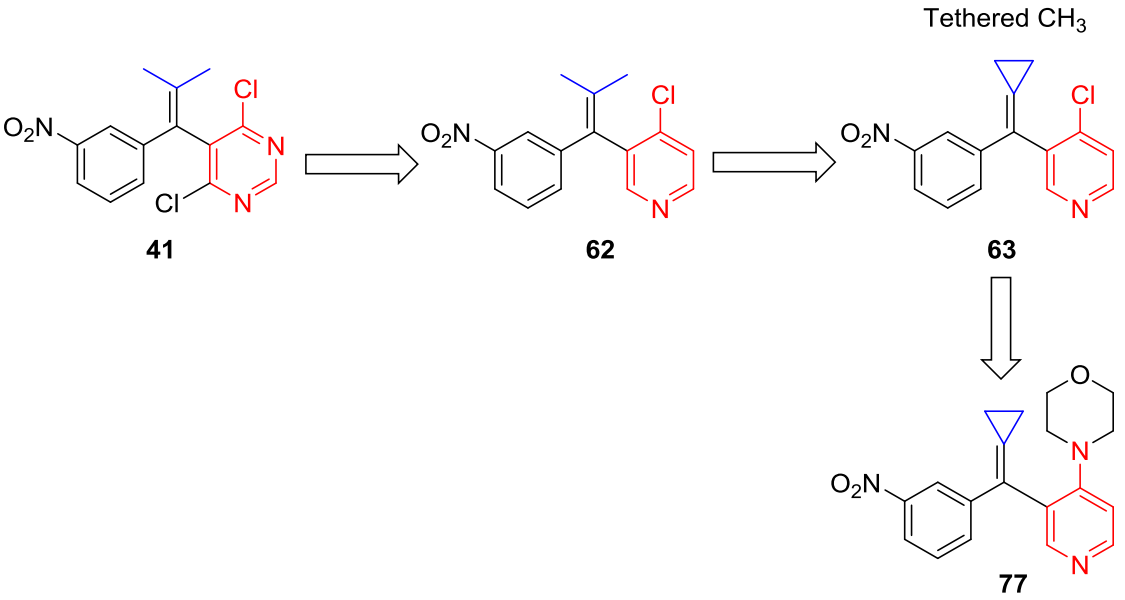
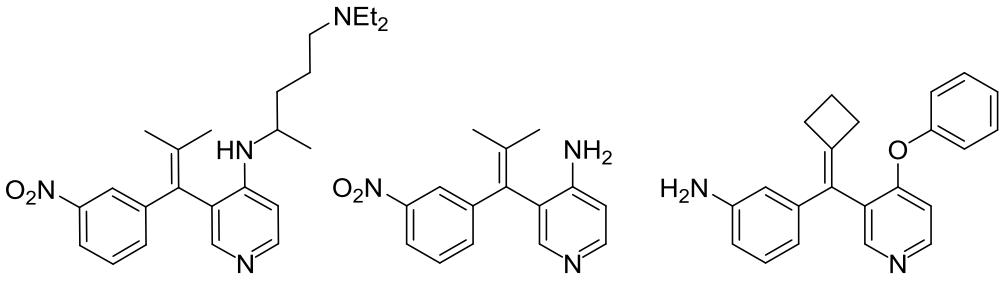


Figure 8.1 Modifications of the AAPM series by simplification and decreasing the steric hindrance.

Different conditions were investigated to synthesize the AAPM with the amines **33** or **47**, however, all attempts failed due to the polymerization of **33** and the decomposition of **47** under the reaction conditions. The degradation of **47** was observed during the attempted synthesis of **65** (Figure 8.2) where the amine **66** peak was

observed in the mass analysis. Furthermore, during the synthesis of **72**, the amine **47** is believed to decompose providing the hydrogen source for the reduction reaction to occur giving the NH_2 group in **72** (Figure 8.2).

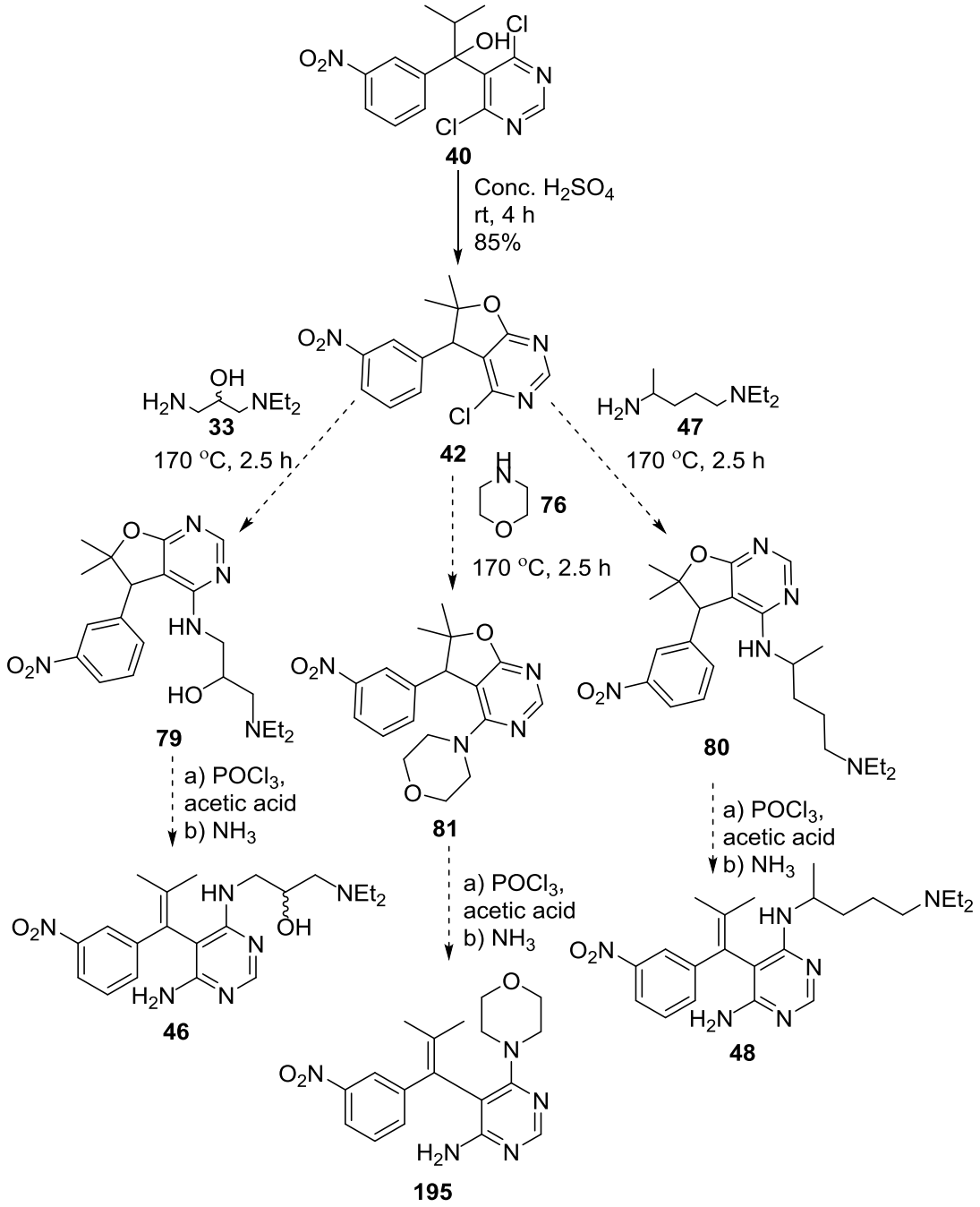


65 **66** **72**

Figure 8.2 The un-synthesized AAPM **65**, the decomposition product **66** and compound **72** that was produced in a reaction with the amine **47**.

8.1.2. Future directions

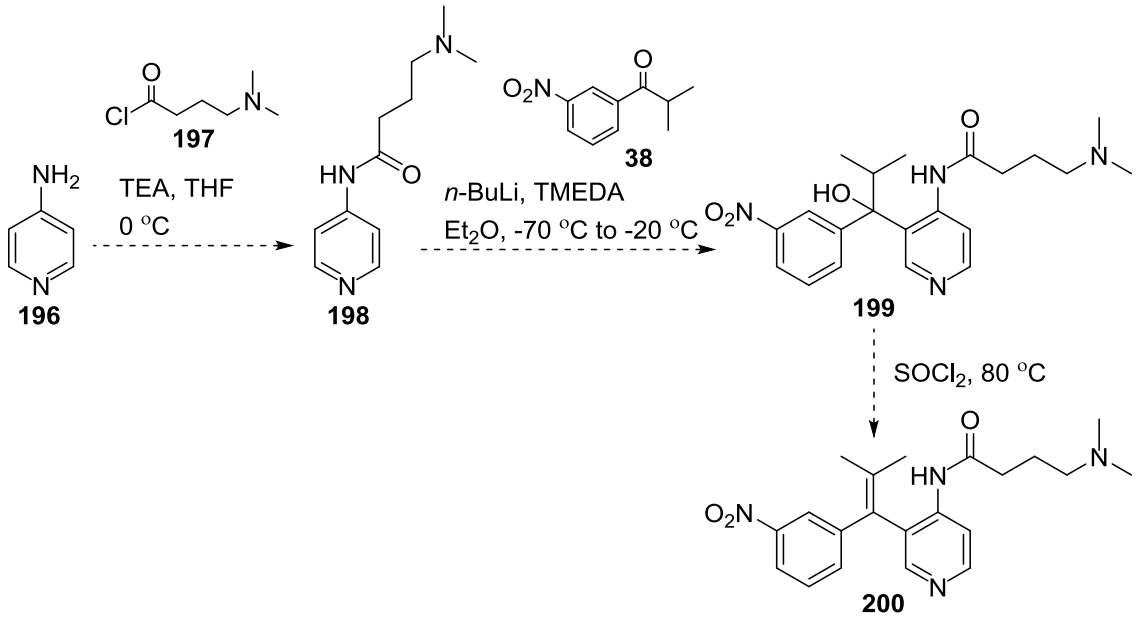
Some alternative access routes to the AAPM derivatives are to be considered. The synthesized new furopyrimidine **42** can be considered for the amination reaction with the original pyrimidine core, as it is less stericaly hindered. The furan ring can then be open using $POCl_3$[160] to give the final aminated products (Scheme 8.1).



Scheme 8.1 Alternative synthesis of the AAPM series using the furopyrimidine **42** intermediate.

Using the pyridine core, an alternative synthesis method to access the AAPM derivatives can be investigated (Scheme 8.2). The suggested synthesis starts with the reaction of the aminopyridine **196** with the amino-acid chloride **197** to afford the amide **198**. Subsequent addition reaction with the ketone **38** using *n*-BuLi and *N,N,N′,N′*-

tetramethylethylenediamine (TMEDA)[261] at -70 ºC would afford the alcohol **199**. Finally, a dehydration reaction would afford the alkene **200** as the AAPM derivative.



Scheme 8.2 An alternative synthetic route for the AAPM derivatives starting from aminopyridine **196**.

8.2. Chapter 3: The CHIKV nsP3

8.2.1. Conclusions

The CHIKV nsP3 protein represents an attractive target for developing selective inhibitors for the virus. This was illustrated by the virtual screening of the NCI set library of 1990 compounds, using the available crystal structure of the macro domain of the nsP3. The hits that achieved higher scores than the co-crystallized ligand (ADP-ribose) were selected for further refining using a second round of docking and scoring. This resulted in 7 hits that scored higher than the co-crystallized ligand. Pose_3 (Figure 8.3) was *in silico* optimized giving pose_3_2, which achieved higher scoring and improved interactions inside the nsP3 macro domain binding pocket. The optimized pose_3_2 also has improved drug-like qualities and fits more to Lipinski's guidelines.[231] The optimized hit was the accessed through three facile chemical synthesis steps. The compounds are still under antiviral evaluations and the accuracy of the model predictions would depend on the biological activity of the compounds.

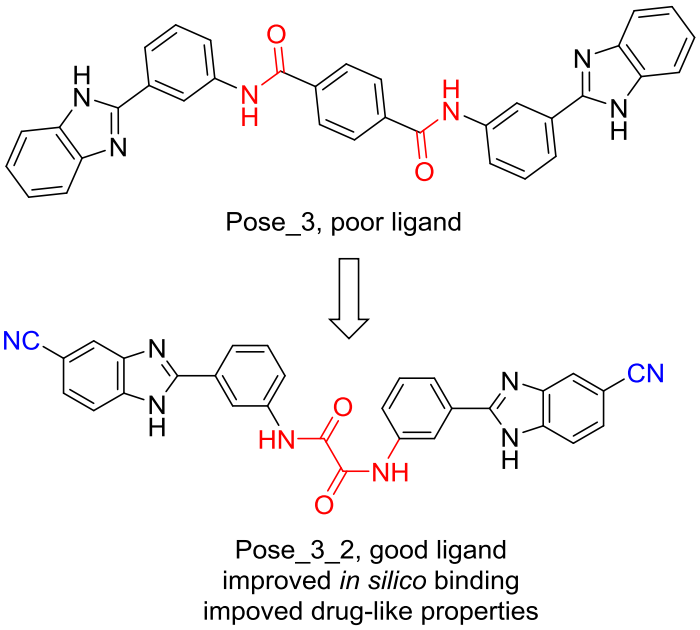


Figure 8.3 The virtual screening hit pose_3 that was *in silico* optimized and synthesized.

8.2.2. Future directions

The parallel virtual screening, *in silico* optimization and *in vitro* evaluation of the obtained hits, would be an applicable structure-based approach for the drug design of selective inhibitors for the CHIKV. Molecular dynamic simulation would also help in optimizing the developed computer model, assisting in obtaining more reliable hits.

Based on the *in vitro* activity of the compounds, the retro-synthesis analyses can be investigated for the active hits, allowing the generation of more analogues which would represent good templates for the discovery of a lead compound, and developing the medicinal chemistry against the CHIKV.

8.3. Chapter 4: Mycophenolic acid analogues

8.3.1. Conclusions

A series of benzolactone-acid conjugates **114-117**, benzolactone-tetrazole conjugates **126**, **127**, **130**, **131** and isatin-acid conjugates **140-143** was synthesized as

mycophenolic acid analogues as inhibitors for the CHIKV (Figure 8.4). Two side products were obtained; the β-lactam ring structure and the open β-lactam product. All the synthesized analogues have drug-like qualities. The access to the isatin-tetrazole conjugates was attempted using the cycloaddition reaction with NaN_3, however, no compounds were obtained due to the reactivity of the isatin core which is believed to interfere with the tetrazole ring formation reaction.

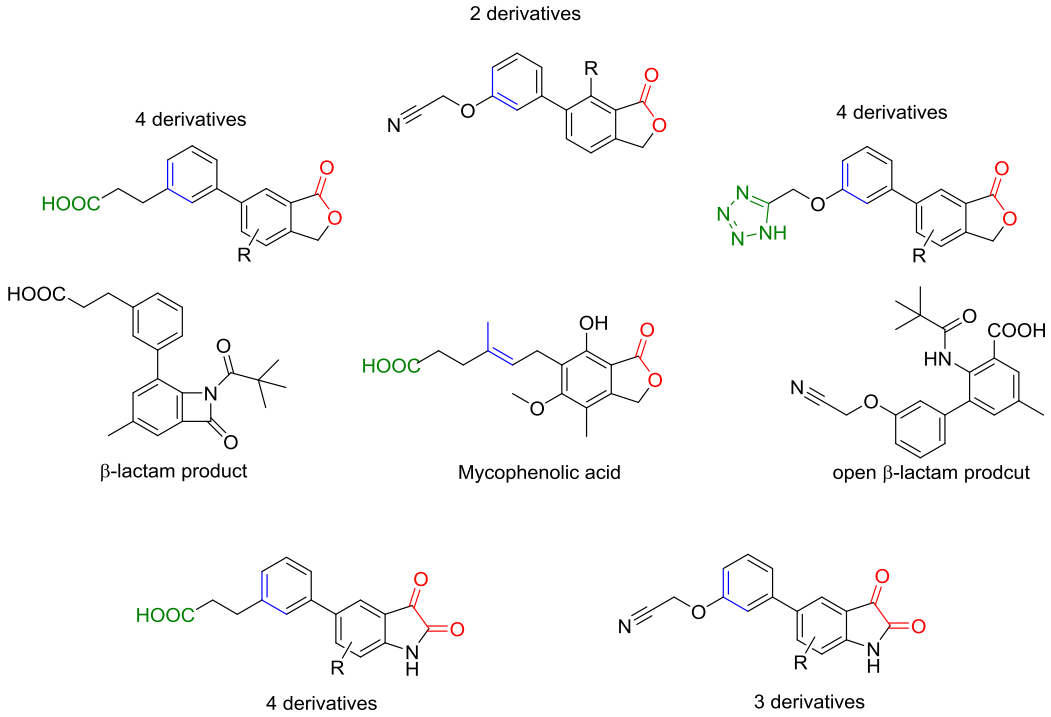


Figure 8.4 Mycophenolic acid derivatives, the coloured features represent the main moieties used in the analogues design.

8.3.2. Future directions

The access to the isatin-tetrazole conjugates could be achieved through the Suzuki coupling with potassium 5-((3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)methyl)tetrazol-1-ide **129**. The synthesized conjugates are currently under anti-chikungunya evaluation, and based on the derived structure activity relationship, further medicinal chemistry derivatisation could be achieved such as formation of esters or amides as prodrugs.

The analogues were designed as mycophenolic acid analogues and are assumed to inhibit the IMPDH enzyme. Therefore, it would be interesting to evaluate these analogues as inhibitors for this enzyme in specific enzyme inhibition tests. Furthermore, a number of crystal structures of the IMPDH enzymes with mycophenolic acid are available within the protein data bank. Therefore, the ability of the analogues to bind to the active site of the enzyme could be *in silico* evaluated, giving more impact to the rational drug design of this series of compounds as spicific enzyme inhibitors.

8.4. Chapter 5: The CHIKV nsP2

8.4.1. Conclusions

The CHIKV protease (nsP2) is critical for the viral replication cycle and therefore, represents an attractive target for drug design. The druggable binding pockets within the nsP2 crystal structure were investigated where two hydrophobic pockets were observed, one within the protease C domain and one within the N domain (Figure 8.5). A computer virtual screening model was developed combining two virtual screening algorithms, FRED and Autodock. The model was used to screen a library of protease inhibitors within the two domains and resulted in favoured hit list for each pocket.

Domain C pocket Domain N pocket

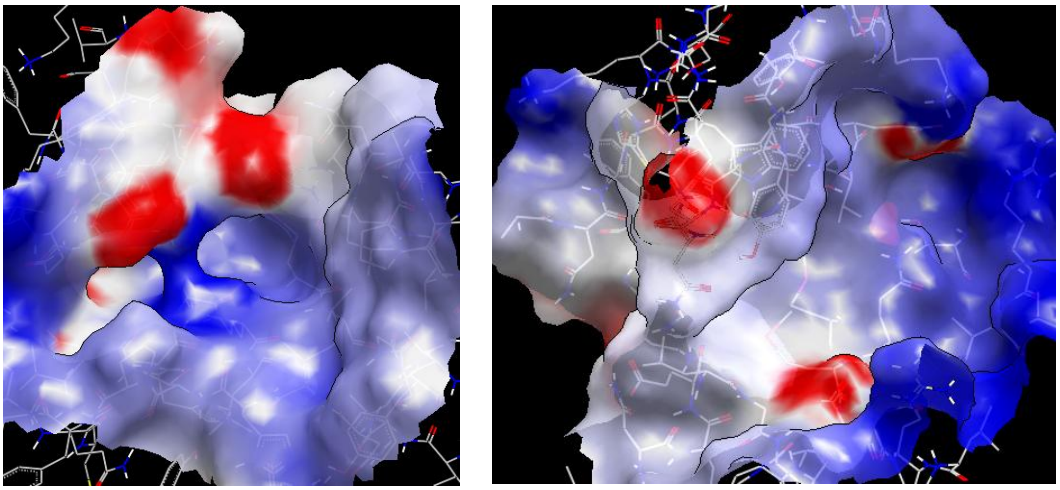


Figure 8.5 The druggable pockets within the CHIKV nsP2.

Within domain C druggable site, Lys1239 (from domain C) and Lys1045 (from domain N) were found to interact with most of the identified ligands, stabilizing them within the U-shaped pocket, blocking the protease function. Within domain N pocket, Trp1084 and Lys1091 were found to be interacting with all of the docked ligands, with the possibility for interaction with the C domain Gln1241. Ligands for this pocket would block the enzyme NTPase/RTPase functions, and might act as allosteric inhibitors for the domain C protease site.

8.4.2. Future directions

The anti-CHIKV evaluation of the identified hits is still ongoing. The results of these assays would be used evaluate the prediction accuracy of the model. The ongoing testing is based on whole cell assays and the ability of the compounds to inhibit the viral induced cytopathic effect and viral replication. The CHIKV nsP2 enzyme is not available as a testing kit for evaluating the direct binding with the model hits, and this would need further investigations and collaborations.

Based on the derived model prediction accuracy, it could be used for screening more compound libraries, with the *in vitro* evaluation of the identified hits. Molecular dynamic simulation studies on the developed model should reveal more insights for the precise molecular interactions with the bound ligands, allowing improved structure-based drug design approach.

8.5. Chapter 6: The CHIKV envelope proteins

8.5.1. Conclusions

The binding sites within the CHIKV envelope proteins (mature and immature forms) were identified with two binding sites found common in the two enzyme forms. A virtual screening computer model combining two successive docking algorithms was developed for these proteins. This model was used for the screening of two chemical libraries within the identified sites in the two enzyme forms. The top 5 hits in each pocket were selected as promising *in silico* inhibitors. The green pocket (Figure 8.6)

represents a surface cavity that lies between the E1 domain II and E2 β–ribbon that connects E2 domain A to E2 domain C, it also extends downwards as a channel between E1 domain II and E2 domain A. The blue pocket (Figure 8.6) is a narrow channel extending just behind the fusion loop and surrounded by both E2 domains A and B. These pockets make contact with residues from E1 and E2 and therefore, ligands for these pockets can affect the relative movement of E1 and E2 domains in the pre- and post-fusion states. Furthermore, the blue pocket makes contact with the fusion loop residues, and therefore, designing antagonists for this pocket would be an applicable strategy to block the fusion function.[262]

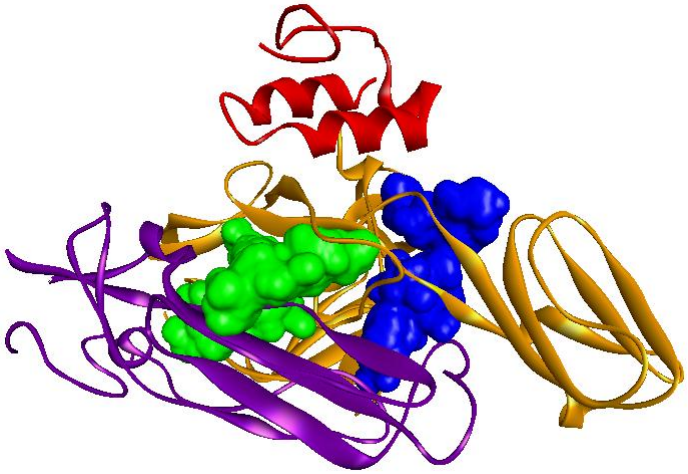


Figure 8.6 Possible druggable cavities within the CHIKV envelope proteins (parts of the whole protein), the blue and the green pockets. E1 is coloured in violet, E2 is coloured in orange and E3 is shown in red.

8.5.2. Future directions

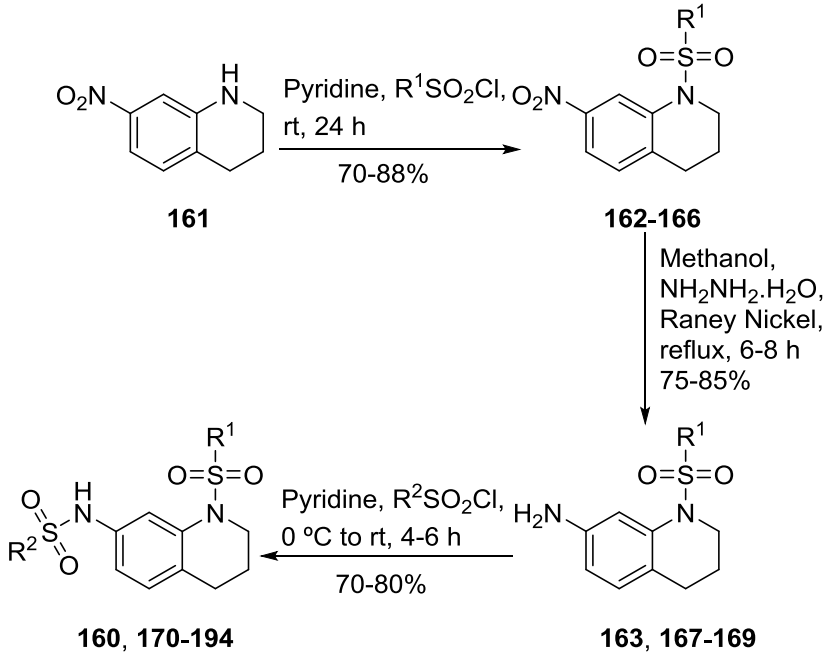
The anti-CHIKV activity of the compounds is to be investigated to evaluate the prediction accuracy of the model. To get more insights into the accuracy of the model, molecular dynamic simulations are to be developed for these models. This developed structure-based computer model can be further used for screening larger database libraries such as the FDA approved drugs libraries that would facilitate the identification of lead compounds against the CHIKV.

Based on the structural similarities of the CHIKV envelope proteins with the other members of *alphaviruses*, this developed computer model as well as the discovered hits could be also evaluated against other viruses.

8.6. Chapter 7: New Leads for African Trypanosomiasis

8.6.1. Conclusions

A bis-sulfonamide lead compound was discovered during the high throughput screening (HTS) of the WEHI 2003 library of 87,926 compounds that were tested against *Trypanosoma brucei*. [256] A series of 26 bis-sulfonamides were synthesized through three facile synthetic steps (Scheme 8.3) and were evaluated against *Trypanosoma brucei* and the cytotoxicity profiles of the compounds were evaluated using the HEK293 cell line. The lead compound was resynthesized and re-evaluated confirming the activity.



Scheme 8.3 The three facile synthesis steps designed for the synthesis of the bis-sulfonamides **160**, **170-194**.

A structure activity relationship was developed for this series based on the activities, with some of the analogues displaying good activity profile ranging from 2 to 4 μM with acceptable selectivity indices from the medicinal chemistry prospective.

8.6.2. Future directions

To further assess the structure activity relationship, another possible derivatisation is the use of amides instead of the sulfonyl linkages (Figure 8.7) giving a series of compounds with the general formula **A**. Another possibility is the change of the core ring, to investigate the importance of the tetrahydroquinoline ring, for example to the indoline ring, giving a series of compounds with the general formulas **B** and **C,** taking into account the preliminary structure activity relationship results. These changes will give a full picture for the requirements of activity for this class of the bis-sulfonamide based skeleton. More biological evaluation are required to assess the mode of action and the target sites for these molecules.

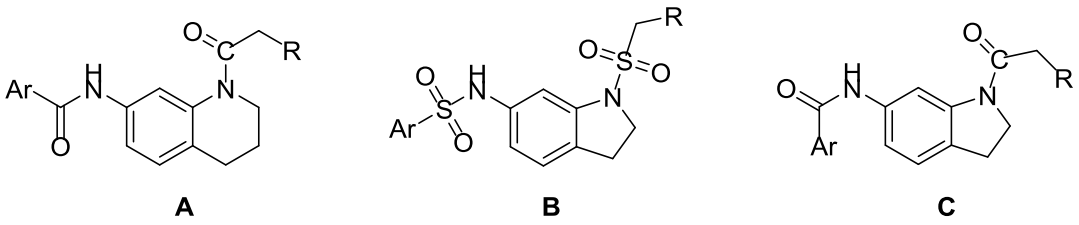


Figure 8.7 Possible derivatisation of the bis-sulfonamide class of compounds.

CHAPTER 9: Experimental

9.1. Chemistry

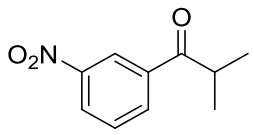
9.1.1. General experimental

Unless otherwise stated, chemicals were purchased from Sigma Aldrich (Australia). 1,2,3,4-Tetrahydro-7-nitroquinoline was purchased from Shanghai Xunxin Chemical Co., Ltd. (China). (3-(Cyanomethoxy)phenyl)boronic acid was purchased from Combi-Blocks, Inc., USA. All ^{1}H, ^{13}C and ^{19}F NMR spectra were recorded at 500, 125 and 282 MHz, respectively on a Varian Inova 500 MHz spectrometer. Chemical shifts (δ) are reported in parts per million relative to TMS using $CDCl_3$ as a solvent unless otherwise noted. The solvents used for ^{13}C NMR were the same used for the ^{1}H NMR unless otherwise noted. Coupling constants (*J*) are reported in Hertz (Hz). Multiplicities are reported as singlet (s), broad (br), doublet (d), doublet of doublets (dd), doublet of doublets of doublets (ddd), doublet of triplet (dt), multiplet (m), pentet (p), sextet (sex) or septet (sp). Electron impact (EI) and electrospray (ES) mass spectra (MS) were recorded on a Shimadzu QP-5000 spectrometer and high resolution (HR) on a VG AutoSpec spectrometer. Electrospray (ESI) mass spectra were recorded on a Micromass Platform LCZ spectrometer and high resolution on a Micromass QTOF2 spectrometer. Ion mass to charge (*m/z*) values are stated with their relative abundances as a percentage in parentheses. Peaks assigned to the molecular ion are denoted by M+ or M- (when using the ESI-). Thin Layer Chromatography (TLC) was performed using Merck Silica Gel F254 aluminium sheets. Column chromatography purifications were performed using flash Silica gel (0.04-0.06 mm, 230 – 400 mesh). Infrared (IR) spectra were recorded on a (Shimadzu FT-IR spectrometer) fitted with a Smart Omni-Sampler germanium crystal accessory. All IR spectra were recorded as neat samples. In the IR spectra, strong, medium and weak peaks are assigned as s, m and w, respectively. Melting points were determined using a Gallenkamp (Griffin) melting point apparatus. Temperatures are expressed in degrees Celsius (^{o}C) and are uncorrected. Microwave reactions were performed using a Discover CEM Focused Microwave Synthesis System in 10 mL closed vessels. HPLC purity check was performed using Waters 1525 HPLC pump, Waters^{Tm} 486 absorbance detector (using the wavelength, 254 nm) and Nova-

Pak® Silica 3.9x150 mm C_{18} column. HPLC solvents used in this work: solvent A was hexane, solvent B was isopropanol containing 0.1% TFA, solvent C was isopropanol and solvent D is ethyl acetate. Retention times in minutes for the HPLC results are expressed as R_t. Reactions performed at -78 °C and -116 °C were cooled using $N_{2\,(l)}$/EtOAc and $N_{2\,(l)}$/EtOH slush baths, respectively. The concentration of *n*-BuLi was determined by titration using phenanthroline and 2-butanol in toluene (ca. 1 M). Known compounds were synthesized either according to the reported procedure (references are provided) or using new methods which were described in detail, and the spectral data collected for these reported intermediates matched those reported in the literature. The assignments of the protons and carbons in the spectra were based on the given nomenclatures, unless otherwise stated.

9.1.2. Experimental for chapter 2

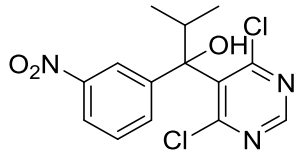
*2-Methyl-1-(3-nitrophenyl)propan-1-one **38***



To conc. H_2SO_4 (25 mL) cooled to -20 ºC, was added dropwise pre-cooled 2-methyl-1-phenylpropan-1-one **37** (20 g, 0.135 mol), and the mixture was stirred for 10 min at -20 ºC. A pre-cooled nitrating mixture of conc. H_2SO_4 (15 mL) and conc. HNO_3 (8.3 g) was added dropwise at a rate that the temperature remained at -10 to -5 ºC. After complete addition, the reaction mixture was stirred for 1 h at 0 ºC. It was then poured onto crushed ice, and extracted with CH_2Cl_2 (2 x 50 mL). The combined organic layers were washed with water (50 mL) and brine (50 mL), dried ($MgSO_4$) and concentrated under reduced pressure. The oily residue was then subjected to flash column chromatography and elution with 40% CH_2Cl_2 in petroleum spirit gave the *meta*-nitro compound **38** (13.5 g, 52%) as a yellow oil; IR (cm^{-1}): υ 1667 (s, C=O), 1528 (s, NO_2), 1355 (s, NO_2). ^{1}H NMR ($CDCl_3$): δ 8.75 (t, *J* = 1.8 Hz, 1H, phenyl_2 H), 8.40 (ddd, *J* = 8.3, 2.3, 1.1 Hz, 1H, phenyl_4), 8.27 (d, *J* = 7.9 Hz, 1H, phenyl_6 H), 7.68 (t, *J* = 8.0 Hz, 1H, phenyl_5 H), 3.57 (sp, *J* = 6.8 Hz, 1H, propanone_2 H), 1.17 (d, *J* = 7.2 Hz, 6H, $(CH_3)_2$). ^{13}C NMR, δ: 202.4 (CO), 149.3 (phenyl C3), 137.7 (phenyl C1), 134.2 (phenyl C6), 130.2 (phenyl C5), 127.4 (phenyl C4), 123.4 (phenyl C2), 36.0 (C2), 19.1 (2 CH_3).

EI-MS *m/z* 193 (M+, 26), 176 (9), 150 (100%), 134 (19), 121 (22), 104 (70). HRMS (ESI) calcd for $C_{10}H_{12}NO_3$ (MH+), 194.0817; found, 194.0823.

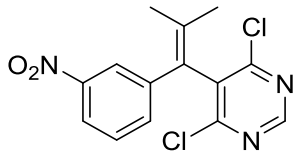
1-(4,6-Dichloropyrimidin-5-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol ***40***



Under a N_2 atmosphere, a fresh solution of LDA was prepared by dropwise addition of *n*-BuLi (1.56 mL of 2 M solution, 3.11 mmol) to a solution of diisopropyl amine (314.7 mg, 3.11 mmol) in a pre-cooled (-78 ºC) dry THF (20 mL), and the mixture was stired for 30 min at -78 ºC. To this solution, was added dropwise, a solution of 4,6-dichloropyrimidine **28** (385.9 mg, 2.59 mmol) in dry THF (10 mL), also pre-cooled to -78 ^{o}C, and the mixture was allowed to stir for 45 min at -78 ^{o}C. The solution was then cooled to -116 ^{o}C and the mixture was allowed to stir for 10 min, after which, a solution of 2-methyl-1-(3-nitrophenyl)propan-1-one **38** (500 mg, 2.59 mmol) in dry THF (10 mL), pre-cooled to -116 ^{o}C, was added dropwise over 10 min period. After complete addition, the resulting solution was allowed to stir for a further 45 min at -116 to -100 ^{o}C. The reaction was then quenched with a saturated solution of NH_4Cl (25 mL). To the resulting biphasic mixture was added H_2O (20 mL) before the organic layer was partitioned from the reaction mixture. The aqueous layer was further extracted with CH_2Cl_2 (3 x 50 mL) and the combined organic layers were washed with 1 M HCl (25 mL), dried ($MgSO_4$) and the solvent concentrated under reduced pressure. The resulting crude oil was subjected to flash column chromatography and elution with 15% ethyl acetate in petroleum spirit gave 1-(4,6-dichloropyrimidin-5-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol **40** (442.8 mg, 50%) as pale yellow crystals, mp: 141-142 ºC; IR (cm^{-1}): υ 3354 (w, OH), 1556 (w, NO_2), 1375 (s, NO_2). ^{1}H NMR ($CDCl_3$): δ 8.63 (s, 1H, pyrimidine_2 H), 8.41 (t, *J* = 1.0 Hz, 1H, phenyl_2), 8.16 (ddd, *J* = 8.4, 2.1, 1.0 Hz, 1H, phenyl_4 H), 7.80 (ddd, *J* = 7.8, 2.0, 0.8 Hz, 1H, phenyl_6 H), 7.52 (t, *J* = 8.1 Hz, 1H, phenyl_5 H), 3.64-3.58 (m, 1H, propanol_2 H), 3.38 (br, OH), 1.11 (d, *J* = 6.6 Hz, 3H, CH_3), 0.93 (d, *J* = 6.6 Hz, 3H, CH_3). ^{13}C NMR, δ: 161.4 (pyrimidine C4, C6), 155.1 (pyrimidine C2), 148.0 (phenyl C3), 145.4 (phenyl C1), 136.3 (pyrimidine C1), 132.5

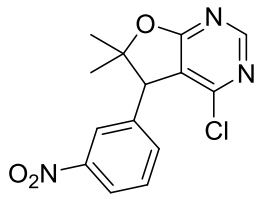
(phenyl C6), 129.1 (phenyl C5), 122.8 (phenyl C4), 121.8 (phenyl C2), 81.7 (propanol C1), 34.9 (propanol C2), 18.1 (CH_3), 17.1 (CH_3). EI-MS *m/z* 342 (M^+, $^{35}Cl_2$, 5), 302 (20), 300 (70), 298 (100%), 284 (15), 282 (20). HRMS (ESI) calcd for $C_{14}H_{14}{}^{35}Cl_2N_3O_3$ (MH+), 342.0407; found, 342.0397.

*4,6-Dichloro-5-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyrimidine **41***



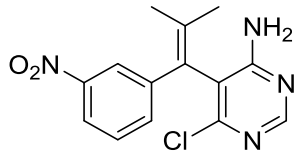
A solution of 1-(4,6-dichloropyrimidin-5-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol **40** (315 mg, 0.963 mmol) in $SOCl_2$ (10 mL) was heated at reflux for 48 h under a N_2 atmosphere. The mixture was then cooled to room temperature, quenched by the dropwise addition of cold H_2O (20 mL), and neutralised with 2 M KOH, and was extraced with CH_2Cl_2 (3 x 50 mL). The combined organic layers were dried ($MgSO_4$), the solvent concentrated under vacuum and the resulting residue was subjected to flash column chromatography (50% CH_2Cl_2 in petroleum spirit) to yield the alkene **41** (220 mg, 70%) as a pale yellow crystaline solid, mp: 93-94 °C; IR (cm^{-1}): υ 2986 (w, C-H aliphatic), 1506 (s, NO_2), 1349 (s, NO_2). ^{1}H NMR ($CDCl_3$): δ 8.74 (s, 1H, pyrimidine_2 H), 8.17 (s, 1H, phenyl_2 H), 8.14 (d, *J* = 8.0 Hz, 1H, phenyl_4 H), 7.60 (d, *J* = 7.5 Hz, 1H, phenyl_4 H), 7.50 (t, *J* = 7.5, 1H, phenyl_5 H), 1.94 (s, 3H, CH_3), 1.73 (s, 3H, CH_3). ^{13}C NMR, δ: 162.2 (pyrimidine C4, C6), 157.1 (pyrimidine C2), 148.3 (phenyl C3), 141.2 (phenyl C1), 140.1 (propene C1), 135.8 (phenyl C6), 134.2 (pyrimidine C5), 129.4 (phenyl C5), 126.0 (propene C2), 124.8 (phenyl C4), 122.6 (phenyl C2), 22.5 (CH_3), 21.8 (CH_3). EI-MS *m/z* 327 (M+ $^{37}Cl_2$, 10), 325 (M+ $^{35}Cl^{37}Cl$, 65), 323 (M+ $^{35}Cl_2$, 100%) 278 (6), 276 (10), 205 (15). HRMS (ESI) calcd for $C_{14}H_{12}{}^{35}Cl_2N_3O_2$ (MH+), 324.0302; found, 324.0298.

4-Chloro-6,6-dimethyl-5-(3-nitrophenyl)-5,6-dihydrofuro[2,3-d]pyrimidine ***42***



A solution of 1-(4,6-dichloropyrimidin-5-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol **40** (50 mg, 0.147 mmol) in conc. H_2SO_4 (5 mL) was stirred for 4 h at room temperature. The mixture was then poured onto crushed ice, and a saturated solution of Na_2CO_3 was added dropwise until no more CO_2 gas evolved. The mixture was then extracted with CH_2Cl_2 (2 x 20 mL) and the combined organic layers were washed sequentially with water (20 mL) and brine (20 mL) and then dried ($MgSO_4$). The solvent was concentrated under reduced pressure and the residue was subjected to flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave **42** (38 mg, 85%) as a pale yellow oil, IR (cm^{-1}): υ 2922 (w, C-H aliphatic), 1522 (s, NO_2), 1350 (s, NO_2), 1238 (m, C-O). ^{1}H NMR ($CDCl_3$), δ: 8.16 (s, 1H, pyrimidine_2 H), 8.22 (d, *J* = 7.5 Hz, 1H, phenyl_4 H), 7.90 (s, 1H, phenyl_2 H), 7.59 (t, *J* = 7.9 Hz, 1H, phenyl_5 H), 7.29 (d, *J* = 7.6 Hz, 1H, phenyl_6 H), 4.55 (s, 1H, C_5 H), 1.73 (s, 3H, CH_3), 1.15 (s, 3H, CH_3). ^{13}C NMR, δ: 174.1 (pyrimidine C2), 159.8 (O-C-N), 157.8 (phenyl C3), 148.3 (pyrimidine C4), 138.9 (phenyl C1), 134.1 (phenyl C6), 130.2 (phenyl C5), 123.5 (phenyl C4), 123.4 (phenyl C2), 119.1 (CCl-C), 90.5 (CH(CH_3)$_2$), 55.2 (C5), 29.9 (CH_3), 24.9 (CH_3) ppm. EI-MS *m/z* 305 (M+ ^{35}Cl, 50), 288 (100%), 270 (30), 248 (20), 230 (17). HRMS (ESI) calcd for $C_{14}H_{13}{}^{35}ClN_3O_3$ (MH+), 306.0645; found, 306.0648.

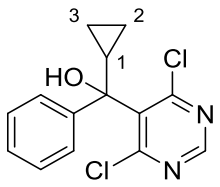
4-Amino-6-chloro-5-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyrimidine ***43***



A suspension of the alkene **41** (100 mg, 0.306 mmol) in NH_4OH (25%, 2 mL) was added to a 10 mL microwave reaction vessel, sealed and heated to 120 °C in a microwave reactor with a maximum pressure of 15 Bar using a power of 50 W,

allowing 10 min (ramping) and then held for 50 min. At this time, the pressure regularly dropped below the 15 Bar maximum and the temperature was increased to 130 °C for a further 30 min. After cooling, the resulting suspension was extracted with ethyl acetate (2 x 25 mL), the combined organic layers were washed with brine (25 mL), dried ($MgSO_4$) and concentrated under vacuum before being subjected to flash column chromatography (80% ethyl acetate in petroleum spirit) to give 4-amino-6-chloro-5-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyrimidine **43** (84.5 mg, 90%) as a yellow crystalline solid, mp: 144-145 °C; IR (cm^{-1}): υ 3329 (w, NH_2), 1562 (s, NO_2), 1347 (s, NO_2). ^1H NMR ($CDCl_3$): δ 8.31 (s, 1H, pyrimidine_2 H), 8.13-8.11 (m, 2H, phenyl_2,4 H), 7.57 (d, *J* = 7.5 Hz, 1H, phenyl_6 H), 7.49 (t, *J* = 8.0 Hz, 1H, phenyl_5 H), 5.27 (br, NH_2), 2.17 (s, 3H, CH_3), 1.97 (s, 3H, CH_3). ^{13}C NMR ($CDCl_3$), δ: 162.0 (pyrimidine C6), 159.0 (pyrimidine C4), 157.3 (pyrimidine C2), 148.2 (phenyl C3), 141.1 (phenyl C1), 140.5 (propene C1), 135.3 (phenyl C6), 129.3 (phenyl C5), 125.3 (propene C2), 124.1 (phenyl C4), 122.3 (phenyl C2), 116.7 (pyrimidine C5), 22.3 (CH_3), 21.9 (CH_3). ESI-MS *m/z* 307 (M+ ^{37}Cl), 305 (M+ ^{35}Cl). HRMS (ESI) calcd for $C_{14}H_{14}{}^{35}ClN_4O_2$ (MH+), 305.0799; found, 305.0799.

Cyclopropyl(4,6-dichloropyrimidin-5-yl)(phenyl)methanol ***50****

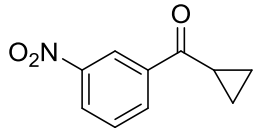


Under a N_2 atmosphere, a fresh solution of LDA was prepared by dropwise addition of *n*-BuLi (1.64 mL of 2.5 M solution, 4.109 mmol) to a solution of diisopropyl amine (415.8 mg, 4.109 mmol) in dry THF (20 mL) pre-cooled to -78 °C, and the mixture was stirred for 30 min at -78 °C. To this solution was added dropwise at -78 °C, a solution of 4,6-dichloropyrimidine **28** (510 mg, 3.424 mmol) in dry THF (10 mL) and the mixture was allowed to stir for 45 min at -78 °C. The solution was then cooled to -116 °C and allowed to stir for 10 min, after which a pre-cooled (-116 °C) solution of cyclopropyl(phenyl)methanone **49** (500 mg, 3.424 mmol) in dry THF (10 mL) was

* Numbers on the structure correspond to the cyclopropyl ring, not related to the compound name. Compound was unstable for HRMS or IR.

added dropwise over a 10 min period. After complete addition, the resulting solution was allowed to stir for a further 45 min at -116 to -100 ^{o}C, then quenched with a saturated solution of NH_4Cl (25 mL). To the resulting biphasic mixture was added H_2O (20 mL) before the organic layer was partitioned from the reaction mixture. The aqueous layer was further extracted with CH_2Cl_2 (3 x 50 mL) and the resulting organic layers were combined and washed with 1 M HCl (25 mL). The organic layer was then dried ($MgSO_4$) and the solvent concentrated under reduced pressure. The resulting crude oil was subjected to flash chromatography and elution with 15% ethyl acetate in petroleum spirit gave the alcohol, cyclopropyl(4,6-dichloropyrimidin-5-yl)(phenyl)methanol **50** (373 mg, 37%) was obtained as a yellow solid, mp: 146-147 °C; ^{1}H NMR ($CDCl_3$), δ: 8.61 (s, 1H, pyrimidine_2 H), 7.38 (d, *J* = 7.0 Hz, 2H, phenyl_2,6 H), 7.35-7.31 (m, 3H, phenyl_3,4,5 H), 2.83 (s, br, OH), 2.06 (p, *J* = 7.0 Hz, 1H, C1_H), 0.90 (q, *J* = 7.0 Hz, 1H, C_3 H), 0.77 (t, *J* = 6.5 Hz, 2H, C_3 H, C_2 H), 0.57 (q, *J* = 6.5 Hz, 1H, C_2 H). ^{13}C NMR, δ: 161.9 (pyrimidine C4, C6), 155.3 (pyrimidine C2), 145.9 (phenyl C1), 138.1 (pyrimidine C5), 128.9 (phenyl C2, C6), 128.3 (phenyl C3, C5), 126.2 (pyrimidine C4), 78.3 (methanol C), 20.1 (C1), 4.1 (C3), 2.3 (C2) ppm. ESI-MS *m/z* 295.1 (M+ ^{35}Cl), 297.0 (M+ ^{37}Cl).

Cyclopropyl(3-nitrophenyl)methanone[263*] ***53***



Cyclopropyl(phenyl)methanone **49** (1.0 g, 6.849 mmol) was added dropwise to a pre-cooled mixture of conc. H_2SO_4 (10 mL) and conc. HNO_3 (10 mL) at -15 ºC, over 10 min. The mixture was then allowed to stir at -10 ºC for 2 h before being poured onto crushed ice, and extracted with ether (2 x 30 mL). The combined organic layers were washed sequentially with water (20 mL), and brine (20 mL) and then dried ($MgSO_4$). The solution was passed through short pad of silica gel, and the eluent was evaporated to leave a pale yellow oil of cyclopropyl(3-nitrophenyl)methanone **53** (915 mg, 70%). IR (cm^{-1}): υ 1662 (s, C=O), 1526 (s, NO_2), 1350 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.84 (s,

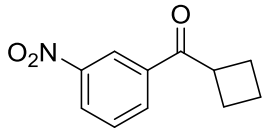
[*] This reference reports only the elemental analysis of **53**.

1H, phenyl_2 H), 8.42 (d, *J* = 8.1 Hz, 1H, phenyl_4 H), 8.33 (d, *J* = 7.8 Hz, 1H, phenyl_6 H), 7.70 (t, *J* = 7.8 Hz, 1H, phenyl_5 H), 2.74-2.69 (m, 1H, CO-C<u>H</u>), 1.32 (t, *J* = 3.7 Hz, 2H, C<u>H</u>$_2$-C<u>H</u>$_2$), 1.16 (p, *J* = 3.7 Hz, 2H, C<u>H</u>$_2$-C<u>H</u>$_2$). ^{13}C NMR, δ: 198.4 (CO), 139.1 (phenyl C3), 133.9 (phenyl C1), 130.6 (phenyl C6), 129.8 (phenyl C5), 126.9 (phenyl C4), 122.9 (phenyl C2), 17.5 (cyclopropyl CH), 12.5 (cyclopropyl CH_2-CH_2) ppm. EI-MS *m/z* 191 (M+, 20), 150 (100%), 104 (70), 77 (75). HRMS (ESI) calcd for $C_{10}H_{10}NO_3$ (MH+), 192.0661; found, 192.0654.

Nitration of Cyclobutyl(phenyl)methanone 56, formation of 57 and 58

Cyclobutyl(phenyl)methanone **56** (5 g, 31.25 mmol) was added to a round bottom two-neck flask fitted with a magnetic stirrer and a thermometer, containing conc. H_2SO_4 (20 mL) and the mixture was cooled to -15 ºC in ice/salt bath. A pre-cooled nitrating mixture of H_2SO_4 (10 mL) and conc. HNO_3 (10 mL) was added dropwise in a rate that the mixture temperature did not exceed 0 ºC. After complete addition, the mixture was stirred for 1 h at 0 - 5 ºC. The mixture was then poured onto crushed ice, and extracted with CH_2Cl_2 (2 x 30 mL). The combined organic layers were then washed with a saturated solution of $NaHCO_3$ (20 mL), water (20 mL), and brine (20 mL) and then dried ($MgSO_4$) and concentrated under reduced pressure. The resulting oil was then subjected to flash column chromatography and elution with 30% ether in petroleum spirit gave:

Cyclobutyl(3-nitrophenyl)methanone **57** (1.73 g, 27%) as a pale yellow oil,

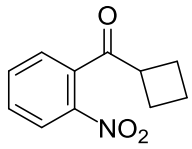


IR (cm^{-1}): υ 2989 (w, C-H aliphatic), 1675 (s, C=O), 1527 (s, NO_2), 1350 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.69 (s, 1H, phenyl_2 H), 8.39 (d, *J* = 7.9 Hz, 1H, phenyl_4 H), 8.23 (d, *J* = 7.5 Hz, 1H, phenyl_6 H), 7.67 (t, *J* = 7.5 Hz, 1H, phenyl_5 H), 4.04 (p, *J* = 8.3 Hz, 1H, CO-CH), 2.46-2.33 (m, 4 H, cyclobutyl), 2.19-2.10 (m, 1H, cyclobutyl), 1.99-1.93 (m, 1H, cyclobutyl). ^{13}C NMR, δ: 198.8 (CO), 148.7 (phenyl C3), 137.1 (phenyl C1), 134.1 (phenyl C6), 130.1 (phenyl C5), 127.4 (phenyl C4), 123.4 (phenyl C2), 42.5

(CO-CH), 25.2 (CH_2-CH_2-CH_2), 18.3 (CH_2-CH_2-CH_2) ppm. EI-MS *m/z* 205 (M+, 5), 188 (10), 150 (100%). HRMS (ESI) calcd for $C_{11}H_{12}NO_3$ (MH+), 206.0817; found, 206.0810.

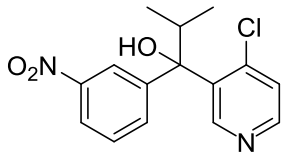
Further elution afforded:

Cyclobutyl(2-nitrophenyl)methanone **58** (640 mg, 10%) as a yellow oil.



Yellow oil, IR (cm^{-1}): υ 2990 (w, C-H aliphatic), 1675 (s, C=O), 1528 (s, NO_2), 1350 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.09 (d, *J* = 8.1 Hz, 1H, phenyl_3 H), 7.70 (t, *J* = 7.6 Hz, 1H, phenyl_5 H), 7.59 (t, *J* = 8.0 Hz, 1H, phenyl_4 H), 7.38 (d, *J* = 7.6 Hz, 1H, phenyl_6 H), 3.66 (p, *J* = 8.5 Hz, 1H, CO-CH), 2.45-2.38 (m, 2H, cyclobutyl 2H), 2.23-2.17 (m, 2H, cyclobutyl 2H), 2.05-1.93 (m, 2H, cyclobutyl 2H). ^{13}C NMR, δ: 204.1 (CO), 146.3, (phenyl C2), 137.3 (phenyl C1), 134.3 (phenyl C5), 130.7 (phenyl C4), 128.0 (phenyl C6), 124.6 (phenyl C3), 45.6 (CO-C), 25.6 (CH_2-CH_2-CH_2), 18.1 (CH_2-CH_2-CH_2) ppm. EI-MS *m/z* 205 (M+, 5), 188 (5), 169 (7), 150 (100%). HRMS (ESI) calcd for $C_{11}H_{12}NO_3$ (MH+), 206.0817; found, 206.0809.

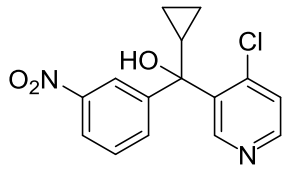
1-(4-Chloropyridin-3-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol ***59***



Under a N_2 atmosphere, diisopropyl amine (400 mg, 3.953 mmol) was added to dry TFH (20 mL) in an oven dried 50 mL flask. The flask was then sealed with a rubber cap, and the flask was degassed by applying vacuum and flushing with N_2 gas. The solution was cooled to -78 ºC before *n*-BuLi (2.47 mL of 1.6 M solution, 3.953 mmol) was added dropwise by a syringe, and the mixture was allowed to stir at -78 ºC for 40 min (Flask A). In a separate oven dried 50 mL flask, 4-chloropyridine.HCl **55** (230 mg, 1.554 mmol) was added and the flask was sealed with a rubber cap and the flask was

flushed with N_2 for 10 min and then sealed; the flask was then cooled to -78 ºC (Flask B). The LDA solution formed in Flask A was then transferred (under N_2) through a double sided needle to Flask B, and the resultant mixture was stirred at -78 ºC for 45 min, where the solution became dark yellow in colour and the 4-chloropyridine.HCl **55** was completely dissolved as can be indicated from the disappearance of the solid. A pre-cooled solution (-78 ºC) of 2-methyl-1-(3-nitrophenyl)propan-1-one **38** (300 mg, 1.554 mmol) in dry THF (10 mL) was added slowly under N_2 to Flask B, using a double sided needle, and the reaction was stirred at -78 ºC for another 45 min. The mixture was allowed to warm up to room temperature (over 12 h) and was then quenched with a mixture of water/THF. The mixture was then acidified with 2 M HCl (20 mL). The aqueous acidic layer was separated and the organic layer was further extracted with 2 M HCl (2 x 25 mL). The combined acidic extracts were washed with ether (25 mL). The acidic extract was then cooled to 0 ºC and conc. ammonia solution was added dropwise to give a slightly basic solution (pH paper). The mixture was then extracted with CH_2Cl_2 (2 x 30 mL), the combined organic layers were dried ($MgSO_4$), and then concentrated under reduced pressure. The residue was subjected to a flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave 1-(4-chloropyridin-3-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol **59** (200 mg, 42%) as a pale yellow solid, mp: 132-133 ºC; IR (cm^{-1}): υ 3180 (w, OH), 1525 (s, NO_2), 1352 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 9.12 (s, 1H, pyridine_2 H), 8.40 (d, *J* = 5.1 Hz, 1H, pyridine_6 H), 8.13 (s, 1H, phenyl_2 H), 8.09 (d, *J* = 7.8 Hz, 1H, phenyl_4 H), 7.69 (d, *J* = 7.8 Hz, 1H, phenyl_6 H), 7.47 (t, *J* = 7.8 Hz, 1H, phenyl_5 H), 7.21 (d, *J* = 5.1 Hz, 1H, pyridine_5 H), 3.25 (s, br, OH), 3.12-3.09 (m, 1H, propanol_2 H), 1.01 (d, *J* = 6.6 Hz, 3H, CH_3), 0.95 (d, *J* = 6.9 Hz, 3H, CH_3). ^{13}C NMR, δ: 149.7 (phenyl C3), 149.6 (pyridine C6), 149.4 (phenyl C1), 147.9 (pyridine C2), 146.2 (pyridine C3), 142.4 (pyridine C4), 137.6 (phenyl C6), 133.3 (phenyl C5), 128.8 (pyridine C5), 126.4 (phenyl C2), 122.2 (phenyl C4), 79.7 (propanol C1), 32.9 (propanol C2), 17.6 (CH_3), 17.4 (CH_3) ppm. EI-MS *m/z* 306 (M+ ^{35}Cl, 5), 263 (100%), 217 (50), 182 (20), 140 (90). HRMS (ESI) calcd for $C_{15}H_{16}{}^{35}ClN_2O_3$ (MH+), 307.0849; found, 307.0838.

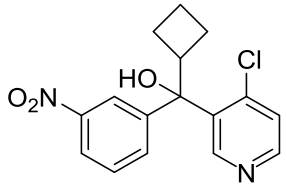
(4-Chloropyridin-3-yl)(cyclopropyl)(3-nitrophenyl)methanol ***60***



Under a N_2 atmosphere, in an oven dried 50 mL flask, a solution of diisopropyl amine (397.4 mg, 3.926 mmol) in dry TFH (20 mL) was prepared, and the flask was then sealed with a rubber cap, and the solution was degassed by applying vacuum and flushing with N_2 gas. The solution was cooled to -78 ºC. *n*-BuLi (2.62 mL of 1.5 M solution, 3.926 mmol) was added dropwise by a syringe, and the mixture was allowed to stir at -78 ºC for 40 min (Flask A). In another oven dried 50 mL flask, 4-chloropyridine.HCl **55** (235 mg, 1.571 mmol) was added and the flask was sealed with a rubber cap and flushed with N_2, and the flask was then cooled to -78 ºC (Flask B). The LDA solution formed in Flask A was then transferred (under N_2) through a double sided needle to Flask B, and the mixture was stirred at -78 ºC for 45 min, where the solution became dark yellow in colour and the 4-chloropyridine.HCl **55** dissolved completely. A pre-cooled solution (-78 ºC) of cyclopropyl(3-nitrophenyl)methanone **53** (300 mg, 1.571 mmol) in dry THF (10 mL) was added slowly under N_2 to the content in flask B, using a double sided needle, and the reaction was stirred at -78 ºC for another 45 min. The mixture was allowed to warm up to room temperature (over 12 h), and was then quenched with a mixture of water/THF. The mixture was then acidified with 2 M HCl. The aqueous acidic layer was separated and the organic layer was further extracted with 2 M HCl (2 x 25 mL). The combined acidic extracts were washed with ether (25 mL). The acidic extract was then cooled to 0 ºC and conc. ammonia solution was added dropwise to give a slightly basic solution (pH paper). The mixture was then extracted with CH_2Cl_2 (2 x 30 mL), dried ($MgSO_4$) and evaporated under reduced pressure. The residue was subjected to a flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave (4-chloropyridin-3-yl)(cyclopropyl)(3-nitrophenyl)methanol **60** (248 mg, 52%) as a yellow solid, mp: 138-139 ºC; IR (cm^{-1}): υ 3100 (w, OH), 1525 (s, NO_2), 1348 (w, NO_2). ^1H NMR ($CDCl_3$), δ: 9.28 (s, 1H, pyridine_2 H), 8.50 (d, *J* = 5.0 Hz, 1H, pyridine_6 H), 8.24 (s, 1H, phenyl_2 H), 8.07 (d, *J* = 8.0 Hz, 1H, phenyl_4 H), 7.57 (d, *J* = 8.0 Hz, 1H, phenyl_6 H), 7.48 (t, *J* = 8.0 Hz, 1H, phenyl_5 H), 7.28 (d, *J* = 5.4 Hz, 1H, pyridine_5 H), 2.98 (s, OH), 1.74 (m,

1H, cyclopropyl_CH), 0.81-0.64 (m, 4H, cyclopropyl_CH_2). ^{13}C NMR, δ: 150.1 (phenyl C3), 149.8 (pyridine C6), 148.4 (pyridine C2), 148.1 (phenyl C1), 143.1 (pyridine C4), 133.9 (phenyl C6), 132.1 (pyridine C3), 129.4 (phenyl C5), 129.1 (phenyl C2), 122.6 (pyridine C5), 121.0 (phenyl C4), 89.4 (methanol C), 20.6 (cyclopropyl CH), 2.4 (CH_2), 2.1 (CH_2) ppm. ESI-MS *m/z* 304.8 (MH+ ^{35}Cl). HRMS (ESI) calcd for $C_{15}H_{14}{}^{35}ClN_2O_3$ (MH+), 305.0693; found, 305.0679.

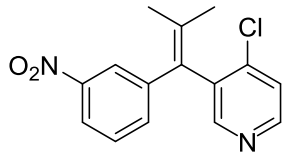
(4-Chloropyridin-3-yl)(cyclobutyl)(3-nitrophenyl)methanol ***61***



Under a N_2 atmosphere, diisopropyl amine (333.2 mg, 3.293 mmol) was added to dry THF (20 mL) in an oven dried 50 mL flask. The flask was then sealed with a rubber cap, and the flask was degassed by applying vacuum and flushing with N_2 gas. The solution was cooled to -78 ºC. *n*-BuLi (2.35 mL of 1.4 M solution, 3.293 mmol) was added dropwise by a syringe, and the mixture was allowed to stir at -78 ºC for 40 min (Flask A). In another oven dried 50 mL flask, 4-chloropyridine.HCl **55** (197 mg, 1.317 mmol) was added and the flask was sealed with a rubber cap and flushed with N_2, and the flask was then cooled to -78 ºC (Flask B). The LDA solution formed in flask A was then transferred (under N_2) through a double sided needle to flask B, and the mixture was stirred at -78 ºC for 45 min, where the solution became dark yellow in colour and the 4-chloropyridine.HCl **55** dissolved completely. A pre-cooled solution (-78 ºC) of cyclobutyl(3-nitrophenyl)methanone **57** (270 mg, 1.317 mmol) in dry THF (10 mL), was added slowly under N_2 to Flask B, using a double sided needle, and the reaction was stirred at -78 ºC for another 45 min. The mixture was allowed to warm up to room temperature (over 12 h), and then was quenched with a mixture of water/THF. The mixture was then acidified with 2 M HCl. The aqueous acidic layer was separated and the organic layer was further extracted with 2 M HCl (2 x 25 mL). The combined acidic extract was washed with ether (25 mL). The acidic extract was then cooled to 0 ºC and conc. ammonia solution was added dropwise to give a slightly basic solution (pH

paper). The mixture was then extracted with CH_2Cl_2 (2 x 30 mL), and dried ($MgSO_4$). The solvent was then evaporated under reduced pressure. The residue was subjected to flash column chromatography (40% ethyl acetate in petroleum spirit) to give (4-chloropyridin-3-yl)(cyclobutyl)(3-nitrophenyl)methanol **61** (209 mg, 50%) as a brown semi solid, IR (cm^{-1}): υ 3094 (w, OH), 1525 (s, NO_2), 1341 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.93 (s, 1H, pyridine_2 H), 8.46 (d, *J* = 5.1 Hz, 1H, pyridine_6 H), 8.15 (s, 1H, phenyl_2 H), 8.09 (d, *J* = 8.0 Hz, 1H, phenyl_6 H), 7.51 (d, *J* = 7.8 Hz, phenyl_4 H), 7.51 (t, *J* = 8.0 Hz, 1H, phenyl_5 H), 7.25 (d, *J* = 4.8 Hz, 1H, pyridine_5 H), 3.44 (p, *J* = 8.6 Hz, 1H, cyclobutyl_CH), 3.33 (s, OH), 2.27-2.18 (m, 2H, cyclobutyl_CH_2), 2.05-1.88 (m, 2H, cyclobutyl_CH_2), 1.85-1.71 (m, 2H, cyclobutyl_CH_2). ^{13}C NMR, δ: 150.3 (phenyl C3), 149.4 (pyridine C6), 148.1 (pyridine C2), 146.7 (phenyl C1), 143.1 (pyridine C3), 136.9 (pyridine C4), 132.2 (phenyl C6), 129.1 (phenyl C5), 126.4 (pyridine C5), 122.2 (phenyl C2), 121.1 (phenyl C4), 42.9 (methanol C), 23.2 (cyclobutyl CH), 22.1 ($\underline{C}H_2$-CH_2-CH_2), 21.1 (CH_2-CH_2-$\underline{C}H_2$), 16.9 (CH_2-$\underline{C}H_2$-CH_2) ppm. EI-MS *m/z* 318 (M+, 7), 301 (5), 263 (100%), 247 (20), 217 (16), 191 (5). HRMS (ESI) calcd for $C_{16}H_{16}ClN_2O_3$ (MH+), 319.0849; found, 319.0842.

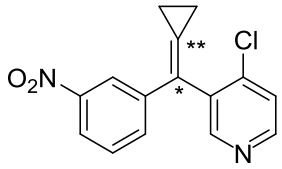
*4-Chloro-3-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyridine **62***



A solution of 1-(4-chloropyridin-3-yl)-2-methyl-1-(3-nitrophenyl)propan-1-ol **59** (100 mg, 0.327 mmol) in conc. H_2SO_4 (20 mL) was stirred under N_2 for 24 h at room temperature. The mixture was then poured onto crushed ice, and neutralized with a saturated solution of Na_2CO_3. The mixture was then extracted with CH_2Cl_2 (2 x 25 mL), the combined organic layers were washed with brine (20 mL), dried ($MgSO_4$), and the solvent was concentrated under reduced pressure. The residue was subjected to a flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave 4-chloro-3-(2-methyl-1-(3-nitrophenyl)prop-1-en-1-yl)pyridine **62** (80 mg, 85%) as a yellow oil, IR (cm^{-1}): υ 2980 (w, C-H aliphatic), 1653 (m, C=C), 1528 (s, NO_2), 1465 (w), 1347 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.49 (s, 1H, pyridine_2 H), 8.42 (d, *J* = 5.4

Hz, 1H, pyridine_6 H), 8.09-8.07 (m, 2H, phenyl_2,6 H), 7.51 (d, *J* = 7.6 Hz, 1H, phenyl_4 H), 7.47 (t, *J* = 7.6 Hz, 1H, phenyl_5 H), 7.35 (d, *J* = 5.4 Hz, 1H, pyridine_5 H), 1.92 (s, 3H, CH_3), 1.72 (s, 3H, CH_3). ^{13}C NMR, δ: 151.9 (phenyl C3), 149.3 (pyridine C6), 148.1 (pyridine C2), 143.6 (phenyl C1), 142.2 (phenyl C6), 138.5 (propene C1), 136.7 (pyridine C3), 135.7 (pyridine C4), 129.0 (phenyl C5), 128.7 (pyridine C5), 124.8 (phenyl C4), 124.5 (phenyl C2), 121.7 (propene C2), 22.5 (CH_3), 21.8 (CH_3) ppm. EI-MS *m/z* 288 (M+ ^{35}Cl, 100%), 273 (5), 241 (10), 227 (16), 204 (8), 191 (10). HRMS (ESI) calcd for $C_{15}H_{14}{}^{35}ClN_2O_2$ (MH+), 289.0744; found, 289.0754.

4-Chloro-3-(cyclopropylidene(3-nitrophenyl)methyl)pyridine ***63***[#]

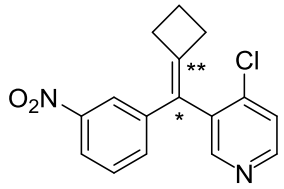


Method A: a solution of 4-chloropyridin-3-yl)(cyclopropyl)(3-nitrophenyl)methanol **60** (100 mg, 0.329 mmol) in conc. H_2SO_4 (10 mL) was stirred under N_2 for 24 h at room temperature. The mixture was then poured onto crushed ice, and neutralized with a saturated solution of Na_2CO_3. The mixture was then extracted with CH_2Cl_2 (2 x 25 mL), the combined organic layers were washed with brine (15 mL), dried ($MgSO_4$), and the solvent was removed under reduced pressure. The residue was subjected to a flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave the alkene **63** (58.3 mg, 62%) as a yellow oil, IR (cm^{-1}): υ 2990 (w, C-H aliphatic), 1676 (w, C=C), 1527 (s, NO_2), 1347 (s, NO_2). ^1H NMR ($CDCl_3$), δ: 8.57 (d, *J* = 5.1 Hz, 1H, pyridine_6 H), 8.50 (s, 1H, pyridine_2 H), 8.17-8.07 (m, 2H, phenyl_2,4 H), 7.49-7.45 (m, 2H, phenyl_6 H, pyridine_5 H), 6.54 (t, *J* = 7.2 Hz, 1H, phenyl_5 H), 3.63 (t, *J* = 6.6 Hz, 2H, C<u>H</u>$_2$-C<u>H</u>$_2$), 2.52 (q, *J* = 6.6 Hz, 2H, C<u>H</u>$_2$-C<u>H</u>$_2$). ^{13}C NMR, δ: 151.7 (phenyl C3), 150.5 (pyridine C6), 148.5 (pyridine C2), 143.8 (phenyl C1), 141.2 (C*), 136.3 (phenyl C6), 132.2 (pyridine C4), 131.4 (phenyl C5), 129.5 (pyridine C3), 124.9 (pyridine C5), 124.0 (phenyl C4), 122.5 (phenyl C2), 121.0 (C**), 43.2 (CH_2), 32.9 (CH_2) ppm. EI-MS *m/z* 286 (M+ ^{35}Cl, 25), 227 (100%), 204 (10), 191 (70), 164 (60). HRMS (ESI) calcd for $C_{15}H_{12}{}^{35}ClN_2O_2$ (MH+), 287.0587; found, 287.0576.

[#] Asterisks on the structures are for peak assignments.

Method B: Under a N_2 atmosphere, (4-chloropyridin-3-yl)(cyclopropyl)(3-nitrophenyl)methanol **60** (100 mg, 0.329 mmol) was added to $SOCl_2$ (15 mL) in a 25 mL round bottom flask. The flask was then heated at 80 ºC under N_2 atmosphere for 24 h. The mixture was cooled to room temperature, and the solution was poured onto crushed ice (50 mL) with vigorous stirring. KOH solution (2 M) was added dropwise until neutral solution obtained (pH paper). The mixture was then extracted with CH_2Cl_2 (2 x 25 mL), washed with water (2 x 20 mL), brine (20 mL) and dried ($MgSO_4$). The organic fraction was then concentrated under reduced pressure and the residue subjected to flash column chromatography (40% ethyl acetate in petroleum spirit) to afford 4-chloro-3-(cyclopropylidene(3-nitrophenyl)methyl)pyridine **63** (70.6 mg, 75%) as a yellow oil. Spectral data for **63** obtained by method B are exactly the same as in method A.

*4-Chloro-3-(cyclobutylidene(3-nitrophenyl)methyl)pyridine **64**[#]*

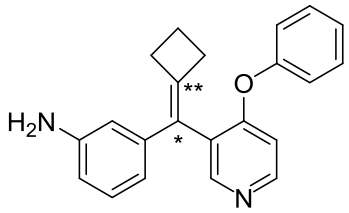


Under a N_2 atmosphere, a solution of (4-chloropyridin-3-yl)(cyclobutyl)(3-nitrophenyl)methanol **61** (90 mg, 0.283 mmol) in $SOCl_2$ (15 mL) was then heated at 80 ºC under N_2 atmosphere for 36 h. The mixture was cooled, and the solution was poured onto crushed ice (50 mL) with vigorous stirring. KOH solution (2 M) was added dropwise to neutral (pH paper). The mixture was then extracted with CH_2Cl_2 (2 x 25 mL), the combined organic extracts were washed sequentially with water (2 x 20 mL), brine (20 mL), and then dried ($MgSO_4$). The solvent was removed under reduced pressure. The resulting residue was subjected to a flash column chromatography and elution with 40% ethyl acetate in petroleum spirit gave 4-chloro-3-(cyclobutylidene(3-nitrophenyl)methyl)pyridine **64** (58.6 mg, 69%) as a brown oil, IR (cm^{-1}): υ 2989 (w, C-H aliphatic), 1657 (m, C=C), 1527 (s, NO_2), 1340 (s, NO_2). ^1H NMR ($CDCl_3$), δ: 8.51 (d, *J* = 5.1 Hz, 1H, pyridine_6 H), 8.49 (s, 1H, pyridine_2 H), 8.04 (d, *J* = 8.1 Hz, 1H,

[#] Asterisks on the structures are for peak assignments.

phenyl_4 H), 8.01 (s, 1H, phenyl_2 H), 7.44 (t, *J* = 8.1 Hz, 1H, phenyl_5 H), 7.40 (d, *J* = 5.6 Hz, 1H, pyridine_5 H), 7.34 (d, *J* = 8.0 Hz, 1H, phenyl_6 H), 3.12 (t, *J* = 7.2 Hz, 2H, cyclobutyl CH_2), 2.68 (t, *J* = 7.2 Hz, 2H, cyclobutyl CH_2), 2.17 (p, *J* = 7.2 Hz, 2H, CH_2-C<u>H</u>$_2$-CH_2). ^{13}C NMR, δ: 152.4 (phenyl C3), 149.9 (pyridine C6), 149.2 (pyridine C2), 148.6 (phenyl C1), 144.1 (C**), 140.3 (C*), 133.9 (phenyl C6), 133.0 (phenyl C5), 129.4 (pyridine C4), 125.7 (pyridine C3), 125.2 (pyridine C5), 121.9 (phenyl C4), 121.5 (phenyl C2), 32.9 (<u>C</u>H_2-CH_2-CH_2), 31.9 (CH_2-CH_2-<u>C</u>H_2), 17.4 (CH_2-<u>C</u>H_2-CH_2) ppm. EI-MS *m/z* 300 (M+ ^{35}Cl, 100%), 285 (5), 253 (20), 218 (20), 164 (22). HRMS (ESI) calcd for $C_{16}H_{14}{}^{35}ClN_2O_2$ (MH+), 301.0744; found, 301.0733.

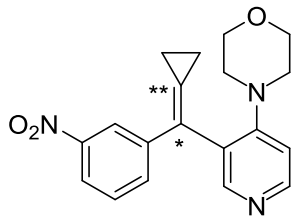
3-(Cyclobutylidene(4-phenoxypyridin-3-yl)methyl)aniline ***72***



4-Chloro-3-(cyclobutylidene(3-nitrophenyl)methyl)pyridine **64 (**100 mg, 0.333 mmol) was added to sealed tube containing phenol (2.0 g, 21.25 mmol) and *N*^{1},*N*^{1}-diethylpentane-1,4-diamine **47** (105 mg, 0.666 mmol). The tube was sealed and heated at 100 ºC for 5 min, 120 ºC for 10 min, 140 ºC for 1 h, and then at 160 ºC for 14 h. The tube was then cooled to room temperature, the mixture was then adsorped onto silica and subjected to a flash column chromatography and elution with 10% methanol in CH_2Cl_2 gave 3-(cyclobutylidene(4-phenoxypyridin-3-yl)methyl)aniline **72** (46.9 mg, 43%) as a colourless oil, IR (cm^{-1}): υ 3256 (w, NH_2), 1452 (s, C-N), 1237 (w, C-O). ^{1}H NMR ($CDCl_3$), δ: 8.41-8.35 (m, 2H, pyridine_2,6 H), 7.33 (t, *J* = 7.8 Hz, 2H, phenoxy_3,5 H), 7.17 (t, *J* = 7.3 Hz, 1H, aniline_5 H), 7.06 (t, *J* = 7.8 Hz, 1H, phenoxy_4 H), 6.90 (d, *J* = 7.8 Hz, 2H, phenoxy_2,6 H), 6.65-6.59 (m, 2H, aniline_4,6 H), 6.50 (d, *J* = 7.8 Hz, 1H, pyridine_5 H), 6.84 (s, 1H, aniline_2 H), 3.55 (br, NH_2), 3.08 (t, *J* = 7.0 Hz, 2H, CH_2-CH_2-C<u>H</u>$_2$), 2.75 (t, *J* = 7.0 Hz, 2H, C<u>H</u>$_2$-CH_2-CH_2), 2.06 (p, *J* = 7.4 Hz, 2H, CH_2-C<u>H</u>$_2$-CH_2).^{13}C NMR, δ: 162.8 (phenoxy C1), 154.9 (pyridine C4), 153.3 (pyridine C2), 150.1 (pyridine C6), 146.4 (aniline C1), 144.6 (aniline C3), 140.9 (C*), 130.2 (phenoxy C3, C5), 129.3 (phenoxy C4), 126.1 (C**), 125.2 (aniline

C5), 120.9 (pyridine C3), 120.8 (phenoxy C2, C6), 118.5 (aniline C4), 114.5 (aniline C2), 113.6 (pyridine C5), 111.8 (aniline C6), 33.1 (CH_2), 32.1 (CH_2), 17.6 (CH_2-$\underline{C}H_2$-CH_2) ppm. EI-MS *m/z* 328 (M+, 100%), 313 (20), 299 (17), 235 (23), 219 (15). HRMS (ESI) calcd for $C_{22}H_{21}N_2O$ (MH+), 329.1654; found, 329.1659.

4-(3-(Cyclopropylidene(3-nitrophenyl)methyl)pyridin-4-yl)morpholine ***77***[#]

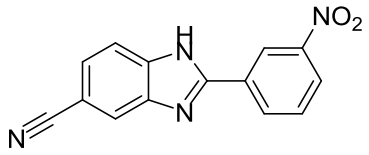


In a 25 mL round bottom flask fitted with a reflux condenser, 4-chloro-3-(cyclopropylidene(3-nitrophenyl)methyl)pyridine **63** (100 mg, 0.349 mmol) was dissolved in morpholine **75** (15 mL), and the mixture was heated at reflux for 48 h. The excess morpholine was then removed by distillation under vacuum to leave a dark brown oily residue, which was subjected to alumina column chromatography using 40% CH_2Cl_2 and 60% ethyl acetate. 4-(3-(Cyclopropylidene(3-nitrophenyl)methyl)pyridin-4-yl)morpholine **77** was isolated as a red oil (76.4 mg, 65%). IR (cm^{-1}): υ 2957 (w, C-H aliphatic), 1525 (s, NO_2), 1457 (m), 1346 (s, NO_2). ^{1}H NMR ($CDCl_3$), δ: 8.56 (d, *J* = 5.3 Hz, 1H, pyridine_6 H), 8.48 (s, 1H, pyridine_2 H), 8.10 (s, 1H, phenyl_2 H), 7.48-7.43 (m, 3H, phenyl_4,6 H, pyridine_5 H), 6.53 (t, *J* = 7.4 Hz, 1H, phenyl_5 H), 3.69 (t, *J* = 4.3 Hz, 4H, morpholine_2,6 H), 2.42 (t, *J* = 7.3 Hz, 2H, cyclopropyl_2H), 2.26 (br, apparent s, 4H, morpholine_3,5 H), 2.08 (ddd, *J* = 14.3, 7.3, 7.3 Hz, 2H, cyclopropyl_2H). ^{13}C NMR, δ: 151.7 (pyridine C6), 150.1 (phenyl C3), 148.4 (pyridine C4), 143.7 (pyridine C2), 141.5 (C*), 134.5 (phenyl C6), 133.7 (phenyl C5), 133.3 (phenyl C1), 131.9 (phenyl C4), 129.3 (phenyl C2), 124.9 (pyridine C3), 122.1 (C**), 120.8 (pyridine C5), 66.8 (morpholine C2, C6), 57.7 (cyclopropyl CH_2), 53.4 (morpholine C3, C5), 27.3 (cyclopropyl CH_2) ppm. ESI-MS *m/z* 336.9 (M-). HRMS (ESI) calcd for $C_{19}H_{19}N_3O_3$ (MH+), 338.1505; found, 338.1506.

[#] Asterisks on the structures are for peak assignments.

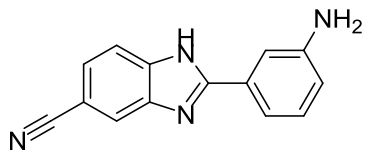
9.1.3. Experimental for chapter 3

*2-(3-Nitrophenyl)-1H-benzo[d]imidazole-5-carbonitrile **84***



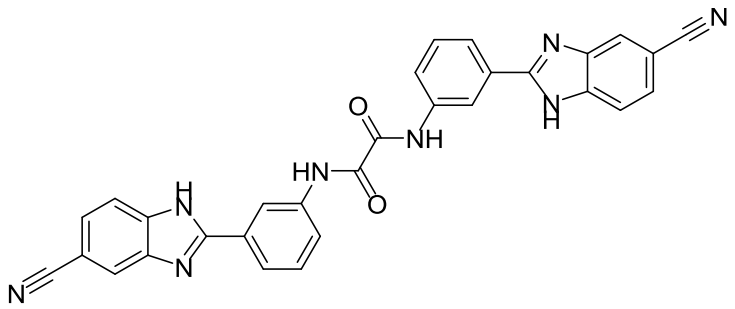
In a 50 mL round bottom flask, 3,4-diaminobenzonitrile **82** (1.0 g, 7.519 mmol) was added to 3-nitrobenzaldehyde **83** (1.13 g, 7.519 mmol), and the mixture was triturated until homogenous mixture obtained. The flask was then fitted with a $CaCl_2$ tube, and heated in an oil bath at 90 ºC with vigorous stirring. The mixture started to melt to a brown oily solution before the mixture solidified into dark brown gummy solid. The oil bath temperature was then increased to 145 ºC where the brown gummy solid started to melt to give dark red oil. After stirring of this red oil for 3 min, the mixture then solidified to give a gummy dark red solid. The flask was then cooled to room temperature and excess ethanol was added and the mixture was heated at 80 ºC for 10 min and filtered hot. The filtrate was concentrated under reduced pressure and the solid residue was recrystallised from a mixture of methanol/1,4-dioxane/water (3:1:1) to give 2-(3-nitrophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile **84** as a yellow crystalline solid (1.38 g, 70%), mp: 239-240 ºC, IR (cm^{-1}): υ 3294 (w, NH), 2224 (m, CN), 1623 (w, C=N), 1517 (s, NO_2), 1348 (s, NO_2). ^{1}H NMR (DMSO-*d6*), δ: 13.8 (br, NH), 9.02 (s, 1H, phenyl_2 H), 8.62 (d, *J* = 7.8 Hz, 1H, phenyl_4 H), 8.37 (d, *J* = 7.8 Hz, 1H, phenyl_6 H), 8.21 (s, 1H, benzimidazole_4 H), 7.89 (t, *J* = 7.8 Hz, 1H, phenyl_5 H), 7.80 (d, *J* = 8.3 Hz, 1H, benzimidazole_6 H), 7.63 (d, *J* = 8.3 Hz, 1H, benzimidazole_7 H). ^{13}C NMR, δ: 152.1 (benzimidazole C2), 148.3 (phenyl C3), 135.4 (benzimdazole C3a), 132.9 (benzimidazole C7a), 132.8 (phenyl C6), 130.8 (phenyl C1), 130.7 (phenyl C5), 126.1 (benzimidazole C6), 124.9 (phenyl C4), 123.9 (benzimidazole C4), 121.3 (phenyl C2), 119.8 (CN), 116.6 (benzimidazole C7), 104.5 (benzimidazole C5) ppm. ESI-MS *m/z* 264.1 (M+). HRMS (ESI) calcd for $C_{14}H_9N_4O_2$ (MH+), 265.0726; found, 265.0723.

2-(3-Aminophenyl)-1H-benzo[d]imidazole-5-carbonitrile ***85***



To a solution of 2-(3-nitrophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile **84** (200 mg, 0758 mmol) in methanol (20 mL) in a 100 mL round bottom flask, was added under a N_2 atmosphere, Raney Nickel (0.1 mol%) with vigorous stirring. Under a N_2 atmosphere, a solution of hydrazine monohydrate (75.6 mg, 1.5 mmol) in methanol (5 mL) was added dropwise over 10 min at room temperature. The mixture was then heated at 80 ºC for 2 h, and filtered hot through a short pad of celite. Methanol and excess hydrazine hydrate were concentrated under reduced pressure to give a brown residue. The residue was then subjected to a flash column chromatography and elution with 10% methanol in CH_2Cl_2 gave 2-(3-aminophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile **85** (141 mg, 80%) as a dark yellow solid, mp: 69-70 ºC, IR (cm^{-1}): υ 3174 (m, NH), 2219 (m, CN), 1616 (m, C=N). ^{1}H NMR (DMSO-*d6*), δ: 8.08 (s, 1H, phenyl_2 H), 7.70 (d, *J* = 8.3 Hz, 1H, phenyl_4 H), 7.56 (d, *J* = 8.3 Hz, 1H, phenyl_6 H), 7.46 (s, 1H, benzimidazole_4 H), 7.32 (d, *J* = 7.2 Hz, 1H, benzimidazole_6 H), 7.19 (d, *J* = 8.0 Hz, 1H, phenyl_5 H), 6.72 (d, *J* = 7.0 Hz, 1H, benzimidazole_7 H), 5.37 (br, NH_2). ^{13}C NMR, δ: 155.4 (benzimidazole C2), 149.2 (phenyl C3), 149.1 (benzimidazole C7a), 140.0 (benzimidazole C3a), 129.8 (phenyl C5), 129.5 (phenyl C1), 125.4 (benzimidazole C6), 120.2 (benzimidazole C4), 116.3 (phenyl C6), 115.9 (CN), 115.5 (phenyl C4), 114.3 (benzimidazole C7), 112.2 (phenyl C2), 103.6 (benzimidazole C5) ppm. ESI-MS *m/z* 233.0 (M-). HRMS (ESI) calcd for $C_{14}H_{11}N_4$ (MH+), 235.0984; found, 235.0993.

N^1,N^2-Bis(3-(5-cyano-1H-benzo[d]imidazol-2-yl)phenyl)oxalamide ***86***



To a solution of 2-(3-aminophenyl)-1*H*-benzo[*d*]imidazole-5-carbonitrile **85** (100 mg, 0.427 mmol) in dry THF (20 mL) In an oven dried 25 mL flask, was added TEA (43 mg, 0.427 mmol) and the mixture was stirred under N_2 at 0 ºC for 30 min. A solution of oxalyl chloride (27.1 mg, 0.213 mmol) in dry THF (5 mL) was added dropwise over 15 min. The mixture was left to warm up to room temperature with stirring for 2 h. Chilled water (10 mL) was added to the reaction mixture with vigorous stirring. The separated solid was filtered, washed with water and methanol, and was recrystallised from DMF to give the bis product **86** (153.8 mg, 69%) as a brown solid, mp: >250 ºC; IR (cm^{-1}): υ 3259 (w, NH), 2224 (m, CN), 1684 (s, C=O). ^{1}H NMR (DMSO-*d6*), δ: 11.08 (s, br, oxalamide NH), 8.85 (s, 2H, phenyl_2 H), 8.18 (s, 2H, benzimidazole_6 H), 7.99-7.96 (m, 4 H, phenyl_4 H, benzimidazole_7 H), 7.78 (d, *J* = 8.2 Hz, 2 H, benzimidazole_6 H), 7.64-7.60 (m, 4H, phenyl_5,6 H). ^{13}C NMR, δ: 158.8 (CO), 153.9 (benzimidazole C2), 141.4 (benzimidazole C7a), 139.9 (phenyl C1), 138.3 (benzimidazole C3a), 129.6 (phenyl C3), 126.5 (benzimidazole C6), 125.9 (benzimidazole C4), 123.2 (phenyl C6), 123.1 (phenyl C5), 120.7 (phenyl C4), 119.9 (CN), 119.5 (phenyl C2), 115.7 (benzimidazole C7), 104.3 (benzimidazole C5) ppm. ESI-MS *m/z* 521.4 (M-H). HRMS (ESI) calcd for $C_{30}H_{17}N_8O_2$ (M-H), 521.1474; found, 521.1453.

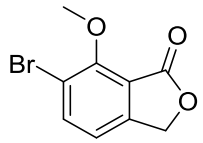
9.1.4. Experimental for chapter 4

General procedure 1: synthesis of isobenzofuran-1(3*H*)-one derivatives 90-92:

Under a N_2 atmosphere, a sealed tube was charged with the appropriate benzoic acid (1 mmol), palladium acetate (22.4 mg, 0.1 mmol), dipotassium phosphate (K_2HPO_4, 522.5 mg, 3 mmol) and dibromomethane (4 mL). The content of the tube was degassed by flushing N_2 through a needle for 10 min. The tube was then sealed, and heated in an oil

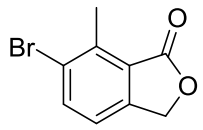
bath at 140 ºC for 36 h. The tube was then allowed to cool to room temperature, the content was diluted with CH_2Cl_2 (20 mL), and then filtered through celite. The filtrate was washed with 1 M HCl (20 mL) and brine (2 x 15 mL), and was dried ($MgSO_4$). The solvent was evaporated under reduce pressure and the residue was either subjected to flash column chromatography (10% methanol in chloroform) or recrystallised from chloroform/hexane mixture (1:3).

*6-Bromo-7-methoxyisobenzofuran-1(3H)-one **90***



Following the general procedure 1using 3-bromo-2-methoxybenzoic acid **87** (231 mg), the benzofuran **90** was isolated after column chromatography, as a white solid (92 mg, 38%), mp: 102-103 ºC; IR (cm^{-1}): υ 1752 (s, C=O), 1288 (m, C-O), 1064 (w, C-O), 1111 (w, C-Br). ^{1}H NMR ($CDCl_3$), δ: 7.84 (d, *J* = 7.9 Hz, 1H, H5), 7.06 (d, *J* = 7.9 Hz, 1H, H4), 5.23 (s, 2H, CH_2), 4.16 (s, 3H, OCH_3). ^{13}C NMR, δ: 167.5 (C=O), 156.3 (C7), 148.3 (C3a), 139.3 (C5), 118.6 (C4), 117.9 (C7a), 117.2 (C6), 68.7 (CH_2), 62.9 (CH_3) ppm. EI-MS *m/z* 242 (M+ ^{79}Br, ^{81}Br, 30), 213 (33), 198 (100%), 183 (15). HRMS (ESI) calcd for $C_9H_8{}^{79}BrO_3$ (MH+), 242.9649; found, 242.9657.

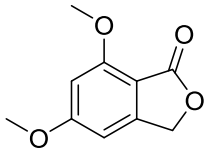
*6-Bromo-7-methylisobenzofuran-1(3H)-one **91***



Following the general procedure 1using 3-bromo-2-methylbenzoic acid **88** (215 mg), the benzofuran **91** was isolated after recrystallisation, as a pale yellow solid (124 mg, 55%), mp: 91-92 ºC; IR (cm^{-1}): υ 1739 (s, C=O), 1358 (w, C-O), 1069 (w, C-Br). ^{1}H NMR ($CDCl_3$), δ: 7.79 (d, *J* = 8.0 Hz, 1H, H5), 7.18 (d, *J* = 8.0 Hz, 1H, H4), 5.20 (s, 2H, CH_2), 2.74 (s, 3H, CH_3). ^{13}C NMR, δ: 170.1 (C=O), 146.1 (C3a), 139.8 (C5), 137.7 (C4), 126.3 (C7), 124.7 (C7a), 120.6 (C6), 68.2 (CH_2), 16.6 (CH_3) ppm. EI-MS *m/z* 226

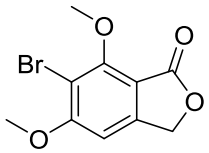
(M+ ^{79}Br, ^{81}Br, 55), 197 (100%), 169 (30). HRMS (ESI) calcd for $C_9H_8{}^{79}BrO_2$ (MH+), 226.9708; found, 226.9705.

5,7-Dimethoxyisobenzofuran-1(3H)-one[264] ***92***



This compound is reported but prepared in this work by a different method, following the general procedure 1 using 2,4-dimethoxybenzoic acid **89** (182 mg). The benzofuran **92** was isolated after column chromatography as a white solid (112.5 mg, 58%). Spectral data of the isolated benzofuran **92** are in agreement with the reported literature.[264]

6-Bromo-5,7-dimethoxyisobenzofuran-1(3H)-one[205*] ***94***

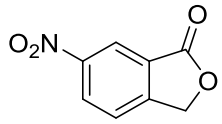


In a round bottom flask (50 mL), 5,7-dimethoxyisobenzofuran-1(3*H*)-one **92** (150 mg, 0.773 mmol) was dissolved in glacial acetic acid (20 mL) with vigorous stirring and the flask was cooled to -5 : 0 ºC using an ice bath. A solution of bromine (136 mg, 0.851 mmol) in glacial acetic acid (10 mL) was added dropwise over 20 min. The reaction was allowed to warm up slowly to room temperature and was then allowed to stir for 4 h. The formed solid was filtered and washed with dilute acetic acid, and was then subjected to column chromatography (60% CH_2Cl_2 in hexane) to afford the benzofuran **94** (176.7 mg, 84%) as a white solid. It could also be recrystallised from 1,4-dioxane, mp: 189-190 ºC; IR (cm^{-1}): υ 1700 (s, C=O), 1316 (w, C-O), 1071 (w, C-O), 1036 (w, C-Br). ^{1}H NMR ($CDCl_3$), δ: 6.46 (s, 1H, H4), 5.07 (s, 2H, CH_2), 4.02 (s, 6H, OCH_3). ^{13}C NMR, δ: 168.4 (C=O), 162.2 (C7), 159.5 (C5), 150.5 (C3a), 107.8 (C7a), 96.2 (C6), 95.1 (C4), 69.3 (CH_2), 57.1 (C_7-O<u>C</u>H_3), 56.6 (C_5-O<u>C</u>H_3) ppm. EI-MS *m/z* 272 (M+

* This reference reports only the elemental analysis and ^{1}H NMR.

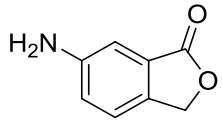
^{79}Br, ^{81}Br, 40), 254 (20), 243 (30), 228 (100%), 213 (20). HRMS (ESI) calcd for $C_{10}H_{10}{}^{79}BrO_4$ (MH+), 272.9762; found, 272.9754.

6-Nitroisobenzofuran-1(3H)-one[208] ***98***



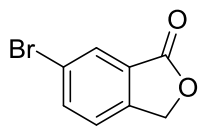
This compound was prepared following the reported procedure.[208] The nitro-phthalide **98** was isolated as a yellow crystalline solid (1.08 g, 81%).

6-Aminoisobenzofuran-1(3H)-one[208] ***99***



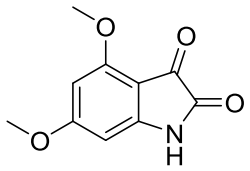
Reduction of the nitro group of the 6-nitroisobenzofuran-1(3*H*)-one **98** to the 6-aminoisobenzofuran-1(3*H*)-one **99** was achieved following the reported reduction procedure developed in our laboratory.[209] by sonicating a suspension of the nitro derivative **98** (172 mg, 0.961 mmol) in a mixture of ethanol (2 mL), acetic acid (2 mL) and water (1 mL) and with iron powder (0.27 g, 5 mmol) for 1 h at 35 ºC. The reaction mixture was filtered to remove the iron residue. The filtrate was partitioned with 2 M KOH, and the basic layer was further extracted with ethyl acetate (3 x 25 mL). The combined organic extracts were washed with brine (2 x 25 mL) and water (3 x 50 mL), dried ($MgSO_4$), and concentrated under reduced pressure. The solid residue was then crystallized from toluene to give the amino derivative **99** in 88% yield (126 mg) as a white solid. The spectral data obtained for this compound **99** using this reduction method are in agreement with the previously reported.[208]

6-Bromoisobenzofuran-1(3H)-one[210] ***100***



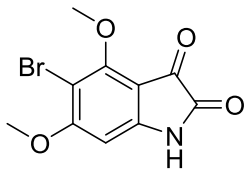
This compound was prepared following the reported procedure.[210] The 6-bromoisobenzofuran-1(3*H*)-one **100** was isolated as a pale yellow solid (106 mg, 74%).

4,6-Dimethoxyindoline-2,3-dione[212] ***102***



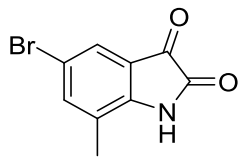
This compound was prepared following the reported procedure.[212] The 4,6-dimethoxyindole-2,3-dione **102** was obtained as a yellow solid (5.2 g, 80%).

5-Bromo-4,6-dimethoxyindoline-2,3-dione ***103***



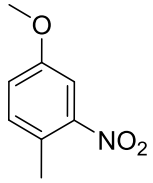
In a 100 mL round bottom flask, a solution of 4,6-dimethoxyindole-2,3-dione **102** (143 mg, 0.5 mmol) in acetic acid (10 mL) was stirred at room temperature until all the isatin was dissolved completely before being cooled to 0 ºC using an ice bath. *N*-Bromosuccinimide (97.9 mg, 0.55 mmol) was added portion wise over 15 min. After addition of the NBS, the reaction mixture started to solidify, and the flask was removed from the ice bath and was allowed to warm to room temperature and then stirred for 8 h. The yellow solid was filtered from the mother liquor to which cold water (10 mL) was added to precipitate the remaining soluble product. The combined solids were washed with saturated $NaHCO_3$ solution (2 x 15 mL) and with 1:1 water ethanol mixture (25 mL) and then vacuum dried. Recrystallisation from 1,4-dioxane yielded the 5-bromo-4,6-dimethoxyindoline-2,3-dione **103** as a yellow solid (118 mg, 60%), mp: >250 ºC; IR (cm^{-1}): υ 3195 (m, NH), 1718 (s, C=O), 1264 (m, C-O). ^{1}H NMR (DMSO-d_6), δ: 11.14 (s, NH), 6.36 (s, H7), 3.99 (s, 3H, C_6-OCH_3), 3.94 (s, 3H, C_4-OCH_3). ^{13}C NMR, δ: 178.4 (C3), 165.5 (C2), 161.7 (C4), 160.4 (C6), 151.4 (C7a), 102.7 (C3a), 91.8 (C7), 86.2 (C5), 58.1 (C4-O<u>C</u>H$_3$), 57.2 (C6-O<u>C</u>H$_3$) ppm. EI-MS *m/z* 287 (M+ ^{79}Br, ^{81}Br, 40), 257 (100%), 213 (7). HRMS (ESI) calcd for $C_{10}H_9N^{79}BrNO_4$ (MH+), 285.9715; found, 285.9719.

5-Bromo-7-methylindoline-2,3-dione[215] ***106***



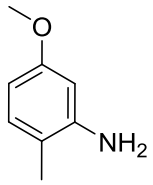
This compound is reported by a different method than the method described here.[215] In this work, 4-bromo-2-methylaniline **104** (6.88 g, 37 mmol) was added to a 500 mL flask and was reacted with chloral hydrate (7.35 g, 44.4 mmol), hydroxylamine hydrochloride (9.25 g, 133 mmol), sodium sulfate (42 g, 295 mmol) in water (250 mL) and 2 M HCl (12.5 mL), by heating the mixture at 55 ºC with vigorous stirring for 14 h. After cooling to room temperature, the aqueous layer was decanted from the formed the gummy residue which was further washed with cold water and dried under vacuum. The flask was then cooled to -5 : 0 ºC using ice/salt bath before adding cold sulphuric acid (22.5 mL). The flask was allowed to warm to room temperature before heating the reaction mixture gradually to 70 ºC where the mixture started to get darker in colour. When the colour change became stable after almost 30 min, the reaction was heated to 80 ºC for 10 min. After cooling to room temperature, the mixture was poured onto crushed ice (150 mL) and stirred for 1 h. The separated solid was then filtered and washed with water (3 x 100 mL) to give a red solid of 5-bromo-7-methylindoline-2,3-dione **106** (6.65 g, 75%) and was sufficiently pure for the next step. The spectral data were consistent with the reported literature.[215]

4-Methoxy-1-methyl-2-nitrobenzene[216] ***108***



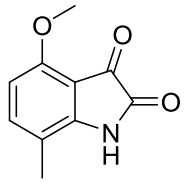
This compound was synthesized following the reported procedure from 4-methyl-3-nitrophenol **107**.[216] The 4-methoxy-1-methyl-2-nitrobenzene **108** was isolated as a white crystalline solid (5.34 g, 80%).

5-Methoxy-2-methylaniline[217] ***109***



The synthesis of this compound is reported by a different method.[217] In this work, it was prepared by dissolving 4-methoxy-1-methyl-2-nitrobenzene **108** (5.1 g, 30.5 mmol) in methanol (35 mL) under a N_2 atmosphere. Raney Nickel (0.1 mol %) was added and the mixture was vigorously stirred at room temperature until the H_2 ceased to evolve. Hydrazine monohydrate (2.0 g, 41 mmol) in methanol (10 mL) was added through a syringe slowly over 10 min. The mixture was then heated at reflux for 4 h. The reaction mixture was filtered hot through celite, and the filtrate was concentrated under reduced pressure. The obtained residue was recrystallised from petroleum spirit to yield 5-methoxy-2-methylaniline **109** as white solid (3.8 g, 91%). Spectral data for this aniline derivative **109** are in agreement with the reported literature.[217]

4-Methoxy-7-methylindoline-2,3-dione[*] ***111***

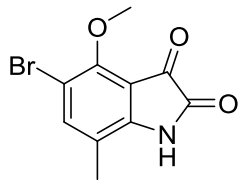


In a 250 mL round bottom flask, 5-methoxy-2-methylaniline **109** (2.0 g, 15 mmol) was dissolved in anhydrous benzene (20 mL), and oxalyl chloride (4.12 g, 2.8 mL, 32.4 mmol) was added dropwise. When the HCl ceased to evolve, the reaction was heated at reflux for 3 h. The benzene and oxalyl chloride were then removed by distillation, and the residue was quickly suspended in dichloroethane (20 mL) and then cooled to 0 ºC in an ice bath. Aluminium chloride (2.1 g, 15.5 mmol) as a suspension in dichloroethane (10 mL) was added portion wise with vigorous stirring and the reaction mixture was allowed to warm to room temperature over 3 h. The reaction mixture was then heated at reflux for 40 min. After cooling to room temperature, the mixture was poured onto

[*] This compound is commercially available but with no synthesis reference.

crushed ice, portioned between 0.1 M HCl and ethyl actetate. The aqueous layer was further extracted with ethyl acetate (2 x 100 mL). The combined organic layers were washed with brine (100 mL), dried ($MgSO_4$) and evaporated under reduced pressure. The resulting orange solid was subjected to a flash column chromatography and elution with 30% ethyl acetate in petroleum spirit yielded 4-methoxy-7-methylindoline-2,3-dione **111** as an orange solid (0.84 g, 30%), mp: 235-236 ºC; IR (cm^{-1}): υ 3195 (m, NH), 1635 (s, C=O), 1250 (m, C-O). ^{1}H NMR (DMSO-d_6), δ: 10.99 (s, NH), 7.36 (d, *J* = 8.7, 1H, H6), 6.61 (d, *J* = 8.7, 1H, H5), 3.83 (s, 3H, OCH_3), 2.09 (s, 3H, CH_3). ^{13}C NMR, δ: 180.9 (C2), 160.1 (C3), 156.4 (C4), 149.1 (C7a), 141.8 (C6), 112.9 (C7), 106.5 (C5), 105.8 (C3a), 55.8 (OCH_3), 14.6 (CH_3) ppm. EI-MS *m/z* 191 (M+, 100%), 163 (50), 149 (10), 134 (7), 121 (7), 106 (93). HRMS (ESI) calcd for $C_{10}H_{10}NNO_3$ (MH+), 192.0661; found, 192.0653.

*5-Bromo-4-methoxy-7-methylindoline-2,3-dione **112***



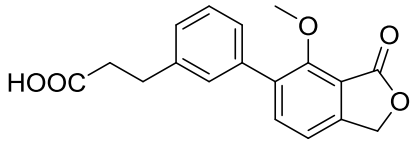
In a 100 mL round bottom flask, 4-methoxy-7-methylindoline-2,3-dione **111** (440 mg, 2.3 mmol) was dissolved in glacial acetic acid (15 mL) and the mixture was stirred until the solid dissolved completely. The flask was then cooled to 0 ºC using an ice bath. A solution of bromine (405 mg, 2.5 mmol) in glacial acetic acid (5 mL) was then added dropwise over 10 min. The flask was allowed to warm to room temperature and was stirred for 2 h. The resulting solid was filtered and cold water (15 mL) added to the filtrate and the resulting solid was filtered. The combined solids were washed with saturated $NaHCO_3$ (2 x 25 mL) and water (2 x 25 mL), and recrystallised from acetic acid to yield 5-bromo-4-methoxy-7-methylindoline-2,3-dione **112** as an orange solid (520.5 mg, 84%), mp: >250 ºC; IR (cm^{-1}): υ 3162 (m, NH), 1696 (s, C=O), 1248 (w, C-O). ^{1}H NMR (DMSO-d_6), δ: 11.20 (s, NH), 7.67 (s, H6), 3.97 (s, 3H, OCH_3), 2.12 (s, 3H, CH_3). ^{13}C NMR, δ: 180.7 (C2), 159.2 (C3), 152.6 (C4), 148.9 (C7a), 142.6 (C6), 117.7 (C7), 110.1 (C3a), 108.1 (C5), 61.8 (OCH_3), 14.6 (CH_3) ppm. EI-MS *m/z* 271

(M+ ^{79}Br, ^{81}Br, 90), 241 (100%), 228 (20), 185 (80). HRMS (ESI) calcd for $C_{10}H_9{}^{79}BrNO_3$ (MH+), 269.9766; found, 269.9773.

General procedure 2: Suzuki coupling of 6-bromobenzolactone derivatives 90, 91, 94 and 100 with 3-(3-boronophenyl)propanoic acid 113:

Under a N_2 atmosphere, a 25 mL flask was charged with the bromoisobenzofuran-1(3*H*)-one (1 equiv), 3-(3-boronophenyl)propanoic acid **113** (1.2 equiv), tetrakis(triphenylphosphine)palladium(0) (5 mol%) and 1,4-dioxane (5 mL). The flask was sealed with a rubber cap. The reaction mixture was degassed by flushing N_2 through a needle for 10 min. The flask was heated in an oil bath at 85 ºC. When the solid dissolved, a N_2 degassed solution of potassium phosphate (3 equiv) in water (3 mL) was added to the reaction mixture and heating at 85 ºC continued for 4-6 h. After cooling to room temperature, the solvent was concentrated under reduced pressure, and 2 M HCl (15 mL) was added to the residue. The suspended solid was either collected by filtration or extracted with ethyl acetate (2 x 25 mL). The combined ethyl acetate extracts were dried ($MgSO_4$) and evaporated under reduced pressure. The obtained residues were dissolved in ethanol and were subjected PLC (10% methanol in CH_2Cl_2, 0.5% acetic acid).

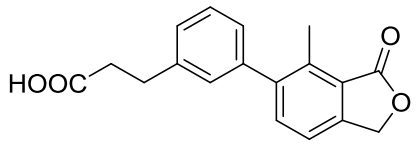
*3-(3-(4-Methoxy-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenyl)propanoic acid **114***



Following the general procedure 2 using 6-bromo-7-methoxyisobenzofuran-1(3*H*)-one **90** (110 mg, 0.453 mmol), 3-(3-boronophenyl)propanoic acid **113** (105.5 mg, 0.544 mmol), tetrakis(triphenylphosphine)palladium(0) (26.2 mg, 0.023 mmol) and potassium phosphate (288.5 mg, 1.359 mmol) and heating for 5 h, the acid **114** was isolated as a white solid (87.5 mg, 62%), mp: 116-117 ºC; IR (cm^{-1}): υ 2989 (m, OH), 1755 (s, C=O), 1718 (s, C=O), 1216 (m, C-O). ^{1}H NMR ($CDCl_3$), δ: 10.54 (s, 1H, COOH), 7.73 (d, 1H, *J* = 7.5 Hz, isobenzofuran_6 H), 7.45 (s, 1H, phenyl_2 H), 7.41 (d, 1H, *J* = 7.7

Hz, phenyl_4 H), 7.40-7.37 (m, 2H, phenyl_5,6 H), 7.30 (d, 1H, *J* = 7.5 Hz, isobenzofuran_7 H), 5.37 (s, 2H, isobenzofuran CH_2), 3.88 (s, 3H, OCH_3), 2.99 (t, 2H, *J* = 6.7 Hz, COOH-C<u>H</u>$_2$-CH_2), 2.68 (t, 2H, *J* = 6.7 Hz, COOH-CH_2-C<u>H</u>$_2$). ^{13}C NMR, δ: 175.6 (COOH), 168.9 (C=O), 156.9 (isobenzofuran C4), 148.4 (phenyl C1), 140.5 (isobenzofuran C7a), 137.8 (phenyl C5), 137.1 (isobenzofuran C6), 135.7 (phenyl C6), 129.5 (phenyl C3), 128.7 (phenyl C2), 127.9 (isobenzofuran C5), 127.5 (phenyl C4), 118.0 (isobenzofuran C3a), 117.1 (isobenzofuran C7), 68.9 (isobenzofuran C1), 62.7 (OCH_3), 47.4 (COOH-<u>C</u>H_2-CH_2), 30.8 (COOH-CH_2-<u>C</u>H_2) ppm. EI-MS *m/z* 312 (M+, 75), 295 (20), 266 (17), 239 (100%), 223 (50), 208 (15), 195 (51). HRMS (ESI) calcd for $C_{18}H_{17}O_5$ (MH+), 313.1076; found, 313.1078. HPLC purity: 97.4%, R_t=8.245 (1% solvent B in solvent A).

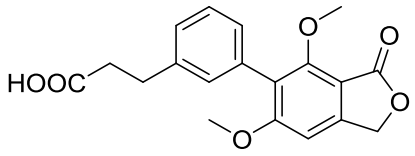
3-(3-(4-Methyl-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenyl)propanoic acid ***115***



Following the general procedure 2 using 6-bromo-7-methylisobenzofuran-1(3*H*)-one **91** (95 mg, 0.418 mmol), 3-(3-boronophenyl)propanoic acid **113** (97.3 mg, 0.502 mmol), tetrakis(triphenylphosphine)palladium(0) (24.2 mg, 0.021 mmol) and potassium phosphate (266.2 mg, 1.254 mmol) and heating for 4 h, the acid **115** was isolated as a white solid (66.8 mg, 54%), mp: 110-111 ºC; IR (cm^{-1}): υ 2989 (m, OH), 1731 (s, C=O), 1696 (s, C=O). ^{1}H NMR ($CDCl_3$), δ: 7.55 (d, 1H, *J* = 7.8 Hz, isobenzofuran_6 H), 7.50 (d, 1H, *J* = 7.3 Hz, phenyl_4 H), 7.37 (t, 1H, *J* = 7.3 Hz, phenyl_5 H), 7.26 (d, 1H, *J* = 7.3 Hz, phenyl_6 H), 7.20 (s, 1H, phenyl_2 H), 7.16 (d, 1H, *J* = 7.3 Hz, isobenzofuran_7 H), 5.28 (s, 2H, isobenzofuran CH_2), 3.02 (t, 2H, *J* = 7.3 Hz, COOH-C<u>H</u>$_2$-CH_2), 2.74 (t, 2H, *J* = 7.3 Hz, COOH-CH_2-C<u>H</u>$_2$), 2.59 (s, 3H, CH_3). ^{13}C NMR, δ: 185.4 (COOH), 171.6 (C=O), 146.3 (phenyl C1), 143.5 (isobenzofuran C7a), 140.5 (phenyl C3), 140.3 (isobenzofuran C4), 137.6 (phenyl C2), 135.7 (isobenzofuran C5), 132.3 (phenyl C5), 129.5 (phenyl C6), 128.7 (isobenzofuran C6), 127.6 (phenyl C4), 123.8 (isobenzofuran C3a), 119.1 (isobenzofuran C7), 68.5 (isobenzofuran C1), 35.7 (COOH-<u>C</u>H_2-CH_2), 30.7 (COOH-CH_2-<u>C</u>H_2), 14.9 (CH_3) ppm. EI-MS *m/z* 296 (M+,

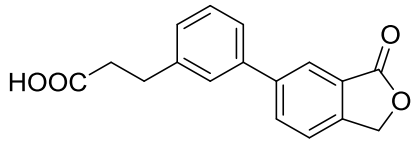
100%), 278 (54), 260 (46), 250 (80), 136 (60), 223 (40). HRMS (ESI) calcd for $C_{18}H_{17}O_4$ (MH+), 297.1127; found, 297.1129. HPLC purity: 97.9%, R_t=5.571 (1% solvent B in solvent A).

3-(3-(4,6-Dimethoxy-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenyl)propanoic acid ***116***



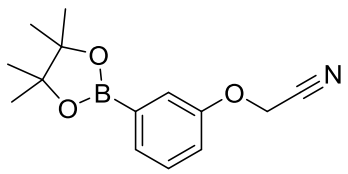
Following the general procedure 2 using 6-bromo-5,7-dimethoxyisobenzofuran-1(3*H*)-one **94** (80 mg, 0.293 mmol), 3-(3-boronophenyl)propanoic acid **113** (68.2 mg, 0.352 mmol), tetrakis(triphenylphosphine)palladium(0) (16.9 mg, 0.015 mmol) and potassium phosphate (186.6 mg, 0.879 mmol) and heating for 6 h, the acid **116** was isolated as a pale yellow oil in (56 mg, 56%), IR (cm^{-1}): υ 2989 (m, OH), 1750 (s, C=O), 1739 (s, C=O), 1208 (w, C-O). ^{1}H NMR ($CDCl_3$), δ: 7.33 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.18 (d, 1H, *J* = 7.9 Hz, phenyl_4 H), 7.12-7.10 (m, 2H, phenyl_2,6 H), 6.50 (s, 1H, isobenzofuran_7 H), 4.98 (s, 2H, isobenzofuran CH_2), 4.03 (s, 3H, isobenzofuran_4 OCH_3), 3.85 (s, 3H, isobenzofuran_6 OCH_3), 2.97 (t, 3H, *J* = 7.6 Hz, COOH-C$\underline{H}_2$-CH_2), 2.66 (t, 3H, *J* = 7.6 Hz, COOH-CH_2-C$\underline{H}_2$). ^{13}C NMR, δ: 177.2 (COOH), 169.0 (C=O), 162.8 (isobenzofuran C6), 159.4 (isobenzofuran C4), 149.3 (phenyl C1), 140.7 (isobenzofuran C7a), 133.8 (phenyl C3), 129.3 (phenyl C2), 128.7 (phenyl C5), 127.7 (phenyl C4), 127.4 (phenyl C6), 117.5 (isobenzofuran C3a), 106.1 (isobenzofuran C5), 95.5 (isobenzofuran C7), 68.3 (isobenzofuran C1), 56.3 (isobenzofuran_6 O$\underline{C}H_3$), 56.2 (isobenzofuran_4 O$\underline{C}H_3$), 35.1 (COOH-$\underline{C}H_2$-CH_2), 30.7 (COOH-CH_2-$\underline{C}H_2$) ppm. EI-MS *m/z* 342 (M+, 100), 326 (25), 296 (26). HRMS (ESI) calcd for $C_{19}H_{18}O_6$ (MH+), 343.1182; found, 343.1190. HPLC purity: 98.7%, R_t=13.234 (5% solvent B in solvent A).

*3-(3-(3-Oxo-1,3-dihydroisobenzofuran-5-yl)phenyl)propanoic acid **117***



Following the general procedure 2 using 6-bromoisobenzofuran-1(3*H*)-one **100** (90 mg, 0.422 mmol), 3-(3-boronophenyl)propanoic acid **113** (98.2 mg, 0.506 mmol), tetrakis(triphenylphosphine)palladium(0) (24.4 mg, 0.021 mmol) and potassium phosphate (268.7 mg, 1.266 mmol) and heating for 4 h, the acid **117** was isolated as a white solid (92.8 mg, 78%), mp: 118-119 ºC; IR (cm^{-1}): υ 2972 (m, OH), 1743 (s, C=O), 1699 (s, C=O). ^{1}H NMR ($CDCl_3$), δ: 8.11 (s, 1H, isobenzofuran_4 H), 7.90 (d, 1H, *J* = 8.0 Hz, phenyl_4 H), 7.55 (d, 1H, *J* = 7.9 Hz, phenyl_6 H), 7.47 (s, 1H, phenyl_2 H), 7.46 (d, 1H, *J* = 6.2 Hz, isobenzofuran_6 H), 7.41 (t, 1H, *J* = 8.0 Hz, phenyl_5 H), 7.27 (d, 1H, *J* = 6.7 Hz, isobenzofuran_7 H), 5.37 (s, 2H, isobenzofuran CH_2), 3.05 (t, 2H, *J* = 7.6 Hz, COOH-C$\underline{H}_2$-CH_2), 2.75 (t, 2H, *J* = 7.6 Hz, COOH-CH_2-C$\underline{H}_2$). ^{13}C NMR, δ: 178.3 (COOH), 171.4 (C=O), 145.5 (phenyl C1), 142.8 (isobenzofuran C7a), 141.4 (phenyl C2), 139.9 (isobenzofuran C5), 133.4 (phenyl C3), 129.6 (phenyl C5), 128.3 (phenyl C6), 127.6 (isobenzofuran C6), 126.7 (isobenzofuran C3a), 125.6 (phenyl C4), 124.3 (isobenzofuran C4), 122.7 (isobenzofuran C7), 69.9 (isobenzofuran C1), 35.6 (COOH-C$\underline{H}_2$-CH_2), 30.8 (COOH-CH_2-C$\underline{H}_2$) ppm. EI-MS *m/z* 282 (M+, 70), 253 (20), 237 (100%), 223 (16). HRMS (ESI) calcd for $C_{17}H_{15}O_4$ (MH+), 283.0970; found, 283.0974. HPLC purity: 99.6%, R_t=6.285 (2% solvent B in solvent A).

*2-(3-(4,4,5,5-Tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetonitrile **119***



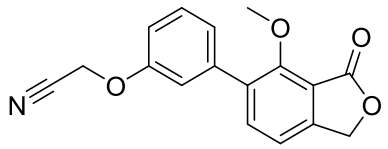
Under a N_2 atmosphere, an oven dried 100 mL flask was charged with 2-(3-iodophenoxy)acetonitrile **124** (1.16 g, 4.479 mmol), bis(pinacolato)diboron **125** (4.55 g, 17.916 mmol), potassium acetate (17.916 mmol, 1.76 g) and [1,1′-

bis(diphenylphosphino)ferrocene]dichloropalladium(II) ($PdCl_2$(dppf)) (32.8 mg, 1 mol%). Dry THF (50 mL) was added to the mixture, and the flask was heated at reflux for 18 h under N_2 atmosphere. The solvent was then evaporated under reduced pressure, ethyl acetate (30 mL) was added and the mixture was sonicated for 10 min and filtered. The filtrate was triturated with $MgSO_4$ and charcoal (0.5 g), and filtered through a short pad of celite. The filtrate was then concentrated under reduced pressure and the resulting residue was recrystallised from a mixture of ethyl acetate/hexane (3:2) to give the boronic ester **119** as a pink solid (1.1 g, 95%), mp: 40-41 ºC; IR (cm^{-1}): υ 2980 (m, C-H aliphatic), 2152 (m, CN), 1128 (w, C-O). ^{1}H NMR ($CDCl_3$), δ: 7.43 (d, 1H, *J* = 7.6 Hz, phenoxy_6 H), 7.39 (s, 1H, phenoxy_2 H), 7.08 (t, 1H, *J* = 7.9 Hz, phenoxy_5 H), 6.96 (d, 1H, *J* = 8.0 Hz, phenoxy_4 H), 4.76 (s, 2H, CH_2), 1.26 (s, 12H, $4CH_3$). ^{13}C NMR, δ: 157.1 (phenoxy C1), 132.6 (phenoxy C3), 131.4 (phenoxy C5), 124.7 (phenoxy C4), 114.8 (CN), 114.6 (phenoxy C2), 94.6 (phenoxy C6), 83.7 (dioxaborolan C4, C5), 53.9 (CH_2), 25.2 ($4CH_3$) ppm. ESI-MS *m/z* 259 (M+). HRMS (ESI) calcd for $C_{14}H_{19}BNO_3$ (MH+), 260.1458; found, 260.1462.

General procedure 3: Suzuki coupling of 6-bromobenzolactone derivatives 90 and 91 with (3-(cyanomethoxy)phenyl)boronic acid 118 or ester 119:

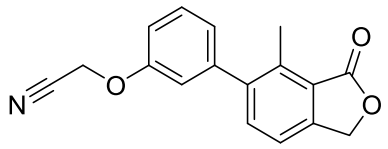
A round bottom 25 mL flask was charged with the appropriate dihydroisobenzofuran derivative (1 equiv), (3-(cyanomethoxy)phenyl)boronic acid **118** (1 equiv) or 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetonitrile **119** (1.2 equiv), tetrakis(triphenylphosphine)palladium(0) (5 mol%), 1,4-dioxane (10 mL) was added and the flask was sealed with a rubber cap. The reaction mixture was degassed by flushing N_2 through a needle for 10 min. The flask was heated in an oil bath at 85 ºC. When all the solid was dissolved, potassium phosphate (3 mmol, 636.8 mg) dissolved in water (5 mL) was added to the reaction *via* a needle. The flask was then heated at 85 ºC for 7-15 h. The solvent was concentrated under reduced pressure, and was added 1M HCl (10 mL), the reaction mixture turned to a white turbid solution which was then extracted with ethyl acetate (2 x 30 mL), washed with brine (2 x 10 mL) and dried ($MgSO_4$). The solvent was evaporated and the residue was subjected to flash column chromatography (40% ethyl acetate in petroleum spirit).

2-(3-(4-Methoxy-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenoxy)acetonitrile ***120***



Following the general procedure 3 using 6-bromo-7-methoxyisobenzofuran-1(3*H*)-one **90** (100 mg, 0.411 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (72 mg, 0.411 mmol), tetrakis(triphenylphosphine)palladium(0) (23.7 mg, 0.021 mmol) and heating for 15 h, the acetonitrile intermediate **120** was isolated as a white solid (60.7 mg, 50%), mp: 103-104 ºC; IR (cm^{-1}): υ 2920 (w, C-H aliphatic), 2165 (m, CN), 1759 (s, C=O), 1212 (w, C-O). ^{1}H NMR ($CDCl_3$), δ: 7.74 (d, 1H, *J* = 7.7 Hz, isobenzofuran_6 H), 7.43 (t, 1H, *J* = 8.0 Hz, phenyl_5 H), 7.25-7.23 (m, 2H, phenyl_4 H, isobenzofuran_7 H), 7.19 (s, 1H, phenyl_2 H), 7.02 (d, 1H, *J* = 8.2 Hz, phenyl_6 H), 5.30 (s, 2H, isobenzofuran CH_2), 4.83 (s, 2H, CN-CH_2), 3.90 (s, 3H, OCH_3). ^{13}C NMR, δ: 168.5 (C=O), 156.7 (phenyl C1), 156.4 (isobenzofuran C4), 148.7 (isobenzofuran C7a), 138.7 (isobenzofuran C6), 137.4 (phenyl C5), 134.7 (isobenzofuran C5), 129.8 (phenyl C4), 124.1 (phenyl C3), 117.9 (phenyl C6), 117.1 (phenyl C2), 116.1 (CN), 115.1 (isobenzofuran C3a), 114.5 (isobenzofuran C7), 68.7 (isobenzofuran C1), 62.7 (CN-$\underline{C}H_2$), 53.7 (OCH_3) ppm. EI-MS *m/z* 295 (M+, 15), 255 (100%), 237 (13), 208 (10), 180 (16). HRMS (ESI) calcd for $C_{17}H_{14}NO_4$ (MH+), 296.0914; found, 296.0913. HPLC purity: 98.3%, R_t=16.676 (1% solvent B in solvent A).

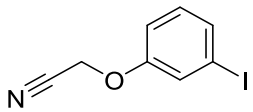
2-(3-(4-Methyl-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenoxy)acetonitrile ***121***



Following the general procedure 3 using 6-bromo-7-methoxyisobenzofuran-1(3*H*)-one **91** (100 mg, 0.358 mmol), 2-(3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy) acetonitrile **119** (111.4 mg, 0.430 mmol), tetrakis(triphenylphosphine)palladium(0) (20.7 mg, 0.018 mmol) and heating for 7 h, the acetonitrile intermediate **121** was isolated as white solid (93.4 mg, 76%), mp: 99-100 ºC; IR (cm^{-1}): υ 2972 (w, C-H

aliphatic), 2169 (m, CN), 1743 (s, C=O). ^{1}H NMR ($CDCl_3$), δ: 7.74 (d, 1H, *J* = 7.7 Hz, isobenzofuran_6 H), 7.43 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.32 (d, 1H, *J* = 7.7 Hz, isobenzofuran_7 H), 7.02 (d, 2H, *J* = 8.0 Hz, phenyl_4,6 H), 6.93 (s, 1H, phenyl_2 H), 5.28 (s, 2H, isobenzofuran CH_2), 4.83 (s, 2H, CN-CH_2), 2.60 (s, 3H, CH_3). ^{13}C NMR, δ: 171.4 (C=O), 156.7 (phenyl C1), 146.7 (isobenzofuran C7a), 142.8 (isobenzofuran C4), 142.1 (isobenzofuran C5), 137.6 (phenyl C3), 135.6 (phenyl C5), 130.1 (isobenzofuran C3a), 124.4 (phenyl C4), 124.1 (isobenzofuran C6), 119.4 (isobenzofuran C7), 116.4 (phenyl C2), 115.3 (CN), 114.1 (phenyl C6), 68.6 (isobenzofuran C1), 53.9 (CN-$\underline{C}H_2$), 15.0 (CH_3) ppm. EI-MS *m/z* 279 (M+, 90), 239 (100%), 223 (15), 195 (12), 165 (24). HRMS (ESI) calcd for $C_{17}H_{14}NO_3$ (MH+), 280.0974; found, 280.0966. HPLC purity: 98.1%, R_t=11.021 (1% solvent B in solvent A).

2-(3-Iodophenoxy)acetonitrile ***124***[265*]



A solution of 3-iodophenol **122** (2.0 g, 9.09 mmol) and NaOH (0.91 g, 22.73 mmol) in dry DMF (25 mL) was stirred vigorously for 30 min until the solution became pale brown. 2-Chloroacetonitrile **123** (1.03 g, 0.86 mL, 13.64 mmol) was then added dropwise through a syringe and the mixture was further stirred for 18 h at room temperature. The resulting brown solution was then poured onto crushed ice (100 mL), stirred for 30 min, and the resulting oil was extracted with ethyl acetate (2 x 25 mL). The combined organic layers were washed with water (2 x 25 mL), brine (2 x 25 mL) and dried ($MgSO_4$). The organic fraction was then triturated with charcoal (0.5 g) and filtered through celite, and evaporated under reduced pressure to afford 2-(3-iodophenoxy)acetonitrile **124** as a yellow oil (1.88 g, 80%), IR (cm^{-1}): υ 2942 (m, C-H aliphatic), 2246 (s, CN), 1203 (w, C-O). ^{1}H NMR ($CDCl_3$), δ: 7.37 (d, 1H, *J* = 7.7 Hz, phenoxy_6 H), 7.29 (s, 1H, phenoxy_2 H), 7.02 (t, 1H, *J* = 7.9 Hz, phenoxy_5 H), 6.90 (d, 1H, *J* = 8.0 Hz, phenoxy_4 H), 4.68 (s, 2H, CH_2). ^{13}C NMR, δ: 157.1 (phenoxy C1), 132.5 (phenoxy C5), 131.6 (phenoxy C4), 124.7 (phenoxy C2), 115.3 (CN), 114.6

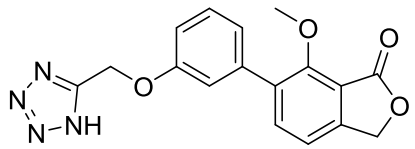
[*] This reference was published in 2013 after the synthesis of **124** in this work, and it only reports the ^{1}H NMR, ^{13}C NMR and LRMS.

(phenoxy C6), 94.9 (phenoxy C3), 54.0 (CH_2) ppm. EI-MS *m/z* 259 (M+, 100%), 219 (27), 203 (5). HRMS (ESI) calcd for C_8H_7INO (MH+), 259.9572; found, 259.9565.

General procedure 4: synthesis of the benzolactone-tetrazole conjugates 126 and 127 using sodium azide method:

Under a N_2 atmosphere, a round bottom flask (25 mL) was charged with the appropriate acetonitrile derivative **120-121** (0.186 mmol), sodium azide (48.4 mg, 0.744 mmol), amonium chloride (39.4 mg, 0.744 mmol), DMF (10 mL) and a catalytic amount of glacial acetic acid (5-10 drops), and the reaction mixture was then heated at reflux for 24 h. DMF was concentrated under reduced pressure and the residue was dried by vacuum. Ethanol was added to the resulting dry residue and the mixture was sonicated for 10 min, filtered and the ethanol solution was concentrated and subjected to PLC (90% CH_2Cl_2, 10% ethanol and 0.1% TFA) to yield the tetrazole conjugate **126** and **127**.

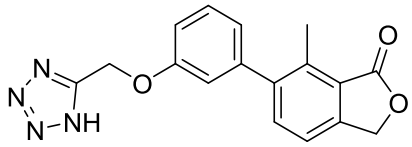
*6-(3-((1H-Tetrazol-5-yl)methoxy)phenyl)-7-methoxyisobenzofuran-1(3H)-one **126***



Following the general procedure 4 using 2-(3-(4-methoxy-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenoxy)acetonitrile **120** (55 mg), the tetrazole conjugate **126** was isolated as a pale yellow solid (21.4 mg, 34%), mp: 130-131 ºC; IR (cm^{-1}): υ 3404 (w, NH), 2907 (w, C-H aliphatic), 1751 (s, C=O), 1660 (m, C=N), 1134 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.72 (d, 1H, *J* = 7.7 Hz, isobenzofuran_5 H), 7.42 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.37 (d, 1H, *J* = 7.7 Hz, isobenzofuran_4 H), 7.22-7.17 (m, 2H, phenyl_2,4 H), 7.10 (d, 1H, *J* = 8.2 Hz, phenyl_6 H), 5.50 (s, 2H, isobenzofuran CH_2), 5.34 (s, 2H, O-CH_2), 3.78 (s, 3H, OCH_3). ^{13}C NMR, δ: 169.7 (C=O), 157.9 (phenyl C3), 156.5 (isobenzofuran C7), 149.6 (tetrazole C), 138.7 (isobenzofuran C3a), 137.7 (isobenzofuran C5), 135.1 (phenyl C5), 129.5 (isobenzofuran C6), 129.4 (phenyl C6), 123.0 (phenyl C1), 117.6 (phenyl C4), 117.5 (phenyl C2), 115.9 (isobenzofuran C7a),

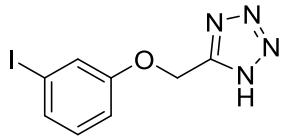
113.9 (isobenzofuran C4), 69.3 (isobenzofuran C3), 61.7 (tetrazole-$\underline{C}H_2$-O), 59.9 (OCH_3) ppm. EI-MS *m/z* 338 (M+, 14), 255 (100%), 239 (10), 211 (12), 195 (11). HRMS (ESI) calcd for $C_{17}H_{15}N_4O_4$ (MH+), 339.1093; found, 339.1107. HPLC purity: 96.3%, R_t=16.339 (5% solvent B in solvent A).

6-(3-((1H-Tetrazol-5-yl)methoxy)phenyl)-7-methylisobenzofuran-1(3H)-one ***127***



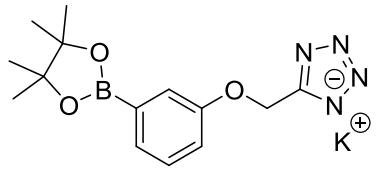
Following the general procedure 4 using 2-(3-(4-methyl-3-oxo-1,3-dihydroisobenzofuran-5-yl)phenoxy)acetonitrile **121** (51.9 mg), the final tetrazole conjugate **127** was isolated as a pale yellow gummy solid (21 mg, 35%), IR (cm^{-1}): υ 3392 (m, NH), 2988 (w, C-H aliphatic), 1741 (s, C=O), 1668 (m, C=N), 1212 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.58 (d, 1H, *J* = 7.7 Hz, isobenzofuran_5 H), 7.48 (d, 1H, *J* = 7.7 Hz, isobenzofuran_4 H), 7.44 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.12 (d, 1H, *J* = 7.9 Hz, phenyl_6 H), 7.03 (s, 1H, phenyl_2 H), 6.99 (d, 1H, *J* = 7.5 Hz, phenyl_4 H), 5.51 (s, 2H, isobenzofuran CH_2), 5.36 (s, 2H, O-CH_2), 2.55 (s, 3H, CH_3). ^{13}C NMR, δ: 170.9 (C=O), 162.4 (phenyl C3), 157.3 (tetrazole C), 147.2 (isobenzofuran C3a), 142.1 (isobenzofuran C7), 141.1 (isobenzofuran C6), 135.7 (phenyl C5), 135.4 (isobenzofuran C5), 129.7 (phenyl C1), 122.9 (isobenzofuran C7a), 122.7 (phenyl C6), 120.8 (isobenzofuran C4), 115.8 (phenyl C2), 114.1 (phenyl C4), 68.5 (isobenzofuran C3), 59.3 (tetrazole-$\underline{C}H_2$-O), 14.4 (CH_3) ppm. EI-MS *m/z* 322 (M+, 50), 239 (75), 223 (26), 195 (60), 165 (100%). HRMS (ESI) calcd for $C_{17}H_{15}N_4O_3$ (MH+), 323.1144; found, 323.1138. HPLC purity: 98.7%, R_t=10.695 (5% solvent B in solvent A).

*5-((3-Iodophenoxy)methyl)-1H-tetrazole **128***



In a round bottom 50 mL flask, a solution of 2-(3-iodophenoxy)acetonitrile **124** (1.6 g, 6.18 mmol), sodium azide (1.6 g, 24.7 mmol) and amonium chloride (1.3 g, 24.7 mmol) in dry DMF (25 mL), containing few drops of glacial acetic acid, was heated at reflux for 14 h under a N_2 atmosphere. The solvent was then concentrated under reduced pressure, and the resulting residue was dried under vacuum. Ethanol (25 mL) was added to the residue and the mixture was sonicated for 10 min, filtered and ethanol was evaporated under reduced pressure resulting in a white fluffy solid of the tetrazole intermediate **128** (1.5 g, 80%), mp: 99-100 ºC ; IR (cm^{-1}): υ 2995 (w, NH), 1589 (m, C=N), 1240 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.46 (s, 1H, phenoxy_2 H), 7.39 (t, 1H, *J* = 7.0 Hz, phenoxy_5 H), 7.09-7.08 (m, 2H, phenoxy_4,6 H), 5.47 (s, 2H, CH_2). ^{13}C NMR, δ: 159.7 (phenoxy C1), 155.3 (tetrazole C), 132.4 (phenoxy C4), 132.2 (phenoxy C5), 125.4 (phenoxy C2), 115.3 (phenoxy C6), 94.9 (phenoxy C3), 61.1 (CH_2) ppm. EI-MS *m/z* 302 (M+, 100%), 220 (80), 203 (10). HRMS (ESI) calcd for $C_8H_8IN_4O$ (MH+), 302.9743; found, 302.9734.

*Potassium 5-((3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)methyl)tetrazol-1-ide **129***



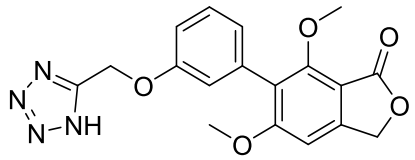
Under a N_2 atmosphere, an oven dried 100 mL flask was charged with 5-((3-iodophenoxy)methyl)-1*H*-tetrazole **128** (960 mg, 3.179 mmol), bis(pinacolato)diboron **125** (3.2 g, 12.716 mmol), potassium acetate (1.25 g, 12.716 mmol) and $PdCl_2$(dppf) (23.3 mg, 1 mol%). Dry THF (30 mL) was added to the mixture, and the flask was heated at reflux for 18 h under a N_2 atmosphere. The solvent was then concentrated under reduced pressure. Ethanol (30 mL) was added to the resulting residue and the

mixture was further heated at reflux for 10 min, filtered hot and concentrated under reduced pressure to afford a white solid of the boronic ester potassium salt **129** (1.1 g, 99%), mp: >250 ºC ; IR (cm^{-1}): υ 2983 (s, C-H aliphatic), 1680 (m, C=N), 1233 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.38 (s, 1H, phenoxy_2 H), 7.28 (d, 1H, *J* = 7.1 Hz, phenoxy_6 H), 7.04-6.99 (m, 2H, phenoxy_4,5 H), 5.25 (s, 2H, CH_2), 1.90 (s, 12H, $4CH_3$). ^{13}C NMR, δ: 178.5 (phenoxy C1), 159.5 (tetrazole C), 158.3 (phenoxy C3), 130.8 (phenoxy C5), 130.2 (phenoxy C4), 124.2 (phenoxy C6), 114.2 (phenoxy C2), 93.6 (dioxaborolan C4, C5), 61.2 (CH_2), 22.6 ($4CH_3$) ppm. ESI-MS *m/z* 301.1 $(M-K)^-$.

General procedure 5: synthesis of the benzolactone-tetrazole conjugates 130 and 131 using the coupling method:

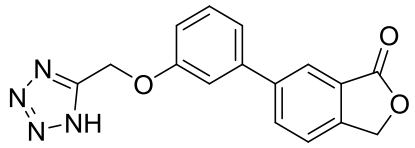
Under a N_2 atmosphere, a mixture of the appropriate bromoisobenzofuran-1(3*H*)-one derivative (1 mmol) and potassium 5-((3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)methyl)tetrazol-1-ide **129** (374 mg, 1.1 mmol), tetrakis(triphenylphosphine)palladium(0) (80.9 mg, 7 mol%) and potassium phosphate (636.8 mg, 3 mmol) was added to a solvent mixture of 1,4-dioxane (25 mL) and water (10 mL) in a 50 mL round bottom flask. The reaction mixture was degassed by flushing N_2 through a needle for 10 min, the flask was then sealed with a rubber cap, heated at 85 ºC for 20-24 h. The solvent was then concentrated under reduced pressure. Saturated $NaHCO_3$ (25 mL) was added to the resulting residues, and the mixture sonicated for 15 min, before the turbid solution was washed with CH_2Cl_2 (25 mL). The aqueous layer was filtered, cooled to 0 ºC using an ice bath, and the filtrate was then neutralized with 2 M HCl. The resulting white solid was collected by vacuum filtration, washed with chilled methanol and dried under vacuum to afford the tetrazoles **130**, **131**.

*6-(3-((1H-Tetrazol-5-yl)methoxy)phenyl)-5,7-dimethoxyisobenzofuran-1(3H)-one **130***



Following the general procedure 5 using 6-bromo-5,7-dimethoxyisobenzofuran-1(3*H*)-one **90** (1 mmol, 273 mg) and heating for 24 h, the final tetrazole conjugate **130** was isolated as a white solid (195 mg, 53%), mp: charring; IR (cm^{-1}): υ 3204 (w, NH), 1740 (s, C=O), 1614 (m, C=N), 1218 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.37 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.05 (d, 1H, *J* = 8.0 Hz, phenyl_6 H), 7.02 (s, 1H, phenyl_2 H), 6.98 (d, 1H, *J* = 7.7 Hz, phenyl_4 H), 6.78 (s, 1H, isobenzofuran_4 H), 5.48 (s, 2H, isobenzofuran CH_2), 5.09 (s, 2H, O-CH_2), 4.04 (s, 3H, isobenzofuran_7 OCH_3), 3.89 (s, 3H, isobenzofuran_5 OCH_3). ^{13}C NMR, δ: 167.9 (C=O), 162.6 (isobenzofuran C5), 158.9 (phenyl C3), 157.4 (isobenzofuran C7), 153.6 (tetrazole C), 149.2 (isobenzofuran C3a), 134.9 (phenyl C5), 129.5 (phenyl C6), 122.8 (phenyl C4), 115.9 (phenyl C10), 115.7 (phenyl C11), 114.1 (isobenzofuran C7a), 104.7 (isobenzofuran C4), 96.2 (isobenzofuran C6), 67.8 (isobenzofuran C3), 59.2 (tetrazole-$\underline{C}H_2$-O), 56.6 (isobenzofuran_7 OCH_3), 56.1 (isobenzofuran_5 OCH_3) ppm. ESI-MS *m/z* 367.1 (M-H). HRMS (ESI) calcd for $C_{18}H_{15}N_4O_5$ ($M\text{-}H^+$), 367.1042; found, 367.1038. HPLC purity: 95.2%, R_t=18.896 (10% solvent B in solvent A).

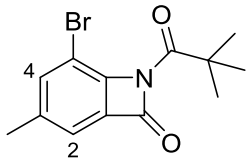
*6-(3-((1H-Tetrazol-5-yl)methoxy)phenyl)isobenzofuran-1(3H)-one **131***



Following the general procedure 5 using 6-bromoisobenzofuran-1(3*H*)-one **91** (213 mg, 1 mmol) and heating for 20 h, the final tetrazole conjugate **131** was isolated as a white solid (181 mg, 59%), mp: charring; IR (cm^{-1}): υ 3195 (w, NH), 1751 (s, C=O), 1653 (m, C=N). ^{1}H NMR (CD_3OD), δ: 8.07 (s, 1H, isobenzofuran_7 H), 7.99 (d, 1H, *J* = 8.0 Hz, isobenzofuran_5 H), 7.67 (d, 1H, *J* = 8.0 Hz, isobenzofuran_4 H), 7.38 (t, 1H, *J* = 7.9 Hz, phenyl_5 H), 7.33 (s, 1H, phenyl_2 H), 7.24 (d, 1H, *J* = 7.5 Hz, phenyl_6 H), 7.09

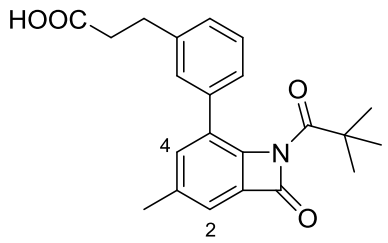
(d, 1H, *J* = 8.0 Hz, phenyl_4 H), 5.42 (s, 2H, isobenzofuran CH_2), 5.37 (s, 2H, O-CH_2). ^{13}C NMR, δ: 171.9 (C=O), 159.3 (phenyl C3), 146.3 (tetrazole C), 142.4 (isobenzofuran C3a), 140.9 (isobenzofuran C6), 133.2 (phenyl C1), 129.7 isobenzofuran C5), 125.9 (phenyl C5), 122.9 (phenyl C6), 122.7 (isobenzofuran C7), 122.6 (isobenzofuran C7a), 119.7 (phenyl C2), 114.4 (isobenzofuran C4), 113.7 (phenyl C4), 69.9 (isobenzofuran C3), 61.1 (tetrazole-$\underline{C}H_2$-O) ppm. ESI-MS *m/z* 307.1 (M-H). HRMS (ESI) calcd for $C_{16}H_{11}N_4O_3$ (M-H^+), 307.0831; found, 307.0826. HPLC purity: 98.8%, R_t=13.175 (5% solvent B in solvent A).

5-Bromo-3-methyl-7-pivaloyl-7-azabicyclo[4.2.0]octa-1,3,5-trien-8-one ***133***



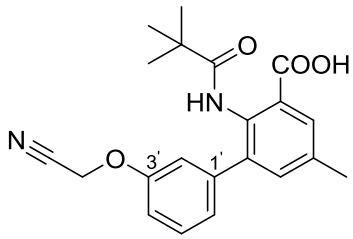
In an oven dried 50 mL flask and under a N_2 atmosphere, 2-amino-3-bromo-5-methylbenzoic acid **132** (750 mg, 3.26 mmol) was dissolved in dry THF (20 mL), and the flask was then sealed with a rubber cap with fitted with a N_2 balloon, and cooled to -78 ºC. *n*-BuLi (4.08 mL of 2 M solution, 8.15 mmol) was added dropwise through a needle and the mixture was stirred for 20 min at -78 ºC before adding pivaloyl chloride (0.48 mL, 3.91 mmol) dropwise. The reaction was then allowed to warm to room temperature and stirred for 3 h. The reaction mixture was then diluted with ethyl acetate (10 mL) and water (10 mL), the organic layer was separated and washed with water (2 x 20 mL), brine (2 x 10 mL) and dried ($MgSO_4$). The combined organic layers were concentrated under reduced pressure to yield the β-lactam derivative **133** as a colourless oil (452 mg, 47%), IR (cm^{-1}): υ 1745 (s, C=O), 1634 (s, C=O), 1465 (w, C-N), 1059 (w, C-Br). ^{1}H NMR ($CDCl_3$), δ: 7.94 (s, 1H, H4), 7.86 (s, 1H, H2), 2.45 (s, 3H, CH_3), 1.42 (s, 9H, pivaloyl 3CH_3). ^{13}C NMR, δ: 168.1 (pivaloyl C=O), 159.7 (C6), 142.3 (C=O), 140.9 (C4), 139.3 (C1), 127.4 (C2), 121.8 (C3), 117.8 (C5), 38.3 ($\underline{C}(CH_3)_3$), 27.6 (pivaloyl 3CH_3), 20.9 (CH_3) ppm. EI-MS *m/z* 295 (M+, 35), 280 (100%), 253 (20), 238 (37). HRMS (ESI) calcd for $C_{13}H_{15}BrNO_2$ (MH+), 296.0286; found, 296.0294.

3-(3-(3-Methyl-8-oxo-7-pivaloyl-7-azabicyclo[4.2.0]octa-1,3,5-trien-5-yl)phenyl)propanoic acid ***137***



Under a N_2 atmosphere, a 25 mL round bottom flask was charged with the β-lactam derivative **133** (100 mg, 0.339 mmol), 3-(3-boronophenyl)propanoic acid **113** (72.3 mg, 0.372 mmol), tetrakis(triphenylphosphine)palladium(0) (19.6 mg, 0.016 mmol) and potassium phosphate (215.9 mg, 1.02 mmol). A mixture of 1,4-dioxane (10 mL) and water (3 mL) was added and the reaction mixture was degassed by flushing N_2 through a needle for 10 min. The flask was then sealed with a rubber cap, heated at 85 ºC for 18 h. The solvent was then concentrated under reduced pressure and the flask was cooled to 0 ºC. Chilled 1 N HCl (10 mL) was added dropwise with stirring, and the resulting solid was extracted with ethyl acetate (2 x 15 mL). The combined organic layers were washed with water (15 mL), brine (15 mL) and dried ($MgSO_4$) and concentrated under reduced pressure. The resulting residue was subjected to a flash column chromatography (50% CH_2Cl_2, 50% ethyl acetate and 1% acetic acid) to afford **137** as a pale yellow oil (74.2 mg, 60%), IR (cm^{-1}): υ 2959 (m, OH), 1696 (s, C=O), 1662 (s, C=O), 1481 (m, C-N). ^{1}H NMR ($CDCl_3$), δ: 8.56 (s, 1H, COOH), 7.77 (s, 1H, H4), 7.32-7.29 (m, 2H, phenyl_5 H, H2), 7.20 (d, 1H, *J* = 7.5 Hz, phenyl_6 H), 7.19 (s, 1H, phenyl_2 H), 7.15 (d, 1H, *J* = 7.7 Hz, phenyl_4 H), 2.97 (t, 2H, *J* = 7.5 Hz, COOH-C$\underline{H}_2$-CH_2), 2.68 (t, 2H, *J* = 7.6 Hz, COOH-CH_2-C$\underline{H}_2$), 2.37 (s, 3H, CH_3), 1.03 (s, 9H, pivaloyl 3CH_3). ^{13}C NMR, δ: 178.4 (pivaloyl C=O), 176.9 (COOH), 171.9 (azabicyclo C=O), 140.4 (phenyl C1), 140.1 (azabicyclo C6), 139.8 (phenyl C3), 136.4 (phenyl C5), 135.8 (azabicyclo C4), 133.1 (azabicyclo C2), 130.9 (azabicyclo C3), 128.8 (azabicyclo C1), 128.7 (phenyl C2), 127.3 (phenyl C6), 126.8 (azabicyclo C5), 124.8 (phenyl C4), 39.3 ($\underline{C}(CH_3)_3$), 35.8 (COOH-$\underline{C}H_2$-CH_2), 30.9 (COOH-CH_2-$\underline{C}H_2$), 27.2 (pivaloyl 3CH_3), 21.0 (CH_3) ppm. EI-MS *m/z* 365 (M+, 60), 350 (70), 320 (50), 299 (25), 277 (100%), 235 (26), 220 (23), 193 (52). HRMS (ESI) calcd for $C_{22}H_{24}NO_4$ (MH+), 366.1705; found, 366.1717. HPLC purity: 98.9%, R_t=27.924 (3% solvent B in solvent A).

*3'-(Cyanomethoxy)-5-methyl-2-pivalamido-[1,1'-biphenyl]-3-carboxylic acid **138***

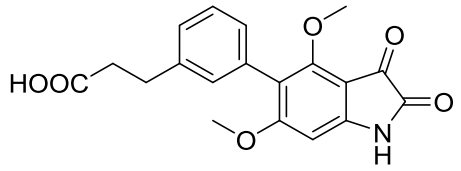


Under a N_2 atmosphere, a 25 mL round bottom flask was charged with the β-lactam derivative **133** (100 mg, 0.339 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (66.0 mg, 0.373 mmol), tetrakis(triphenylphosphine)palladium(0) (19.6 mg, 0.016 mmol) and potassium phosphate (215.9 mg, 1.02 mmol). A mixture of 1,4-dioxane (10 mL) and water (3 mL) was added and the reaction mixture was degassed by flushing N_2 through a needle for 10 min. The flask was then sealed with a rubber cap, heated at 85 ºC for 12 h. The solvent was then concentrated under reduced pressure and 1 N HCl (10 mL) was added dropwise to the resulting residue which was extracted with ethyl acetate (2 x 15 mL). The combined organic layers were washed with water (15 mL), brine (15 mL) and dried ($MgSO_4$). The mixture was then concentrated under reduced pressure, and the resulting residue was subjected to a flash column chromatography and elution with 90% CH_2Cl_2, 10% ethanol and 0.1% TFA yielded **138** as a pale yellow oil (59.4 mg, 48%), IR (cm^{-1}): υ 3328 (m, NH), 2967 (m, OH), 2200 (m, CN), 1646 (w, NH), 1411 (w, C-N). ^{1}H NMR ($CDCl_3$), δ: 8.61 (s, 1H, COOH), 7.53 (s, 1H, phenyl_6 H), 7.32 (t, 1H, *J* = 7.7 Hz, phenyl_5' H), 7.15 (s, 1H, phenyl_4 H), 7.01 (d, 1H, *J* = 7.5 Hz, phenyl_6' H), 6.92-6.90 (m, 2H, phenyl_2', 4'), 6.47 (s, 1H, NH), 4.73 (s, 2H, CH_2), 2.28 (s, 3H, CH_3), 0.97 (s, 9H, pivaloyl 3 CH_3). ^{13}C NMR, δ: 177.7 (pivaloyl C=O), 176.7 (COOH), 173.8 (C3'), 156.4 (C2), 141.7 (C5), 137.8 (C1), 135.7 (C6'), 133.2 (C4), 130.5 (C6), 129.7 (C1'), 129.6 (C6'), 123.4 (C4'), 115.6 (CN), 115.1 (C3), 113.3 (C2'), 53.5 (CH_2), 38.9 ($\underline{C}(CH_3)_3$), 26.9 (pivalamido 3 CH_3), 20.8 (CH_3) ppm. ESI-MS *m/z* 365.0 (M-H). HRMS (ESI) calcd for $C_{21}H_{23}N_2O_4$ (MH+), 367.1658; found, 367.1656. HPLC purity: 99.4%, R_t=17.599 (2% solvent B in solvent A).

General procedure 6: Suzuki coupling of 5-bromoisatin derivatives 103, **106**, **112 and 139 with 3-(3-boronophenyl)propanoic acid 113:**

Under a N_2 atmosphere, a 25 mL flask was charged with the 5-bromoisatin derivative (1 equiv), 3-(3-boronophenyl)propanoic acid **113** (1.2 equiv), tetrakis(triphenylphosphine)palladium(0) (5 mol%) and DME (3 mL) or diglyme (3 mL), was sealed with a rubber cap. The reaction mixture was degassed by flushing N_2 through a needle for 10 min, and was then heated in an oil bath at 85 ºC. When the solid dissolved, a N_2 degassed solution of potassium phosphate (3 equiv) in water (3 mL) was added to the reaction mixture and heating continued at 85 ºC for 4-6 h. After cooling to room temperature, the solvent was concentrated under reduced pressure. A solution of HCl 2 M (15 mL) was added, and the resulting suspended solid was either collected by filtration or extracted with ethyl acetate (2 x 25 mL). The combined organic layers were dried ($MgSO_4$) and evaporated under reduced pressure. The combined residues were then dissolved in ethanol and were subjected to PLC (10% methanol in CH_2Cl_2, 0.5% acetic acid) to yield the isatin-acid conjugates **140**-**143**.

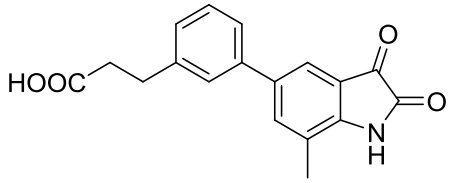
*3-(3-(4,6-Dimethoxy-2,3-dioxoindolin-5-yl)phenyl)propanoic acid **140***



Following the general procedure 6 using 5-bromo-4,6-dimethoxyindoline-2,3-dione **103** (100 mg, 0.349 mmol), 3-(3-boronophenyl)propanoic acid **113** (81.2 mg ,0.419 mmol), tetrakis(triphenylphosphine)palladium(0) (20.2 mg, 0.017 mmol) and potassium phosphate (222.2 mg, 1.047 mmol) and heating for 6 h, the acid **140** was isolated as a yellow solid (49.5 mg, 40%), mp: 175-176 ºC; IR (cm^{-1}): υ 3185 (m, NH), 2978 (m, OH), 1718 (s, C=O), 1696 (s, C=O), 1290 (w, C-O). ^{1}H NMR (DMSO-d_6), δ: 12.09 (s, 1H, COOH), 10.36 (s, NH), 7.31 (t, *J* = 7.5 Hz, 1H, phenyl_5 H), 7.20 (d, *J* = 7.5 Hz, 1H, pheny_4 H), 7.08-7.06 (m, 2H, phenyl_2,6 H), 6.35 (s, 1H, isatin_7 H), 3.97 (s, 3H, isatin_4 OCH_3), 3.84 (s, 3H, isatin_C6 OCH_3), 2.85 (t, *J* = 7.5 Hz, 2H, COOH-C$\underline{H}_2$-CH_2), 2.58 (t, *J* = 7.5 Hz, 2H, COOH-CH_2-C$\underline{H}_2$). ^{13}C NMR, δ: 178.1 (C$\underline{O}$OH), 174.0

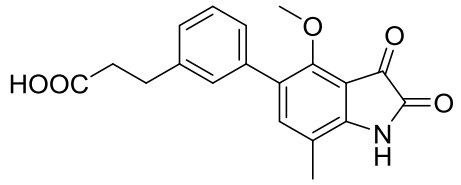
(isatin C2), 166.6 (isatin C3), 161.5 (isatin C6), 160.1 (isatin C4), 149.2 (phenyl C1), 141.0 (isatin C7a), 131.2 (phenyl C3), 130.3 (phenyl C5), 128.3 (phenyl C6), 127.4 (phenyl C2), 120.9 (phenyl C4), 108.2 (isatin C7), 100.6 (isatin C5), 90.1 (isatin C3a), 56.7 (isatin_4 OCH_3), 56.3 (isatin_6 OCH_3), 35.0 (COOH-$\underline{C}H_2$-CH_2), 30.3 (COOH-CH_2-$\underline{C}H_2$) ppm. EI-MS *m/z* 355 (M+, 40), 299 (4), 254 (100%), 191 (45). HRMS (ESI) calcd for $C_{19}H_{18}NO_6$ (MH+), 356.1134; found, 356.1116. HPLC purity: 99.3%, R_t=35.814 (3% solvent B in solvent A).

3-(3-(7-Methyl-2,3-dioxoindolin-5-yl)phenyl)propanoic acid ***141***



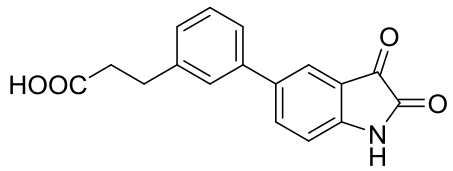
Following the general procedure 6 using 5-bromo-7-methylindoline-2,3-dione **106** (100 mg, 0.416 mmol), 3-(3-boronophenyl)propanoic acid **113** (96.8 mg, 0.499 mmol), tetrakis(triphenylphosphine)palladium(0) (24 mg, 0.021 mmol) and potassium phosphate (264.9 mg, 1.248 mmol) and heating for 4 h, the acid **141** was isolated as a red solid (70 mg, 55%), mp: 238-239 ºC; IR (cm^{-1}): υ 3172 (m, NH), 2989 (m, OH), 1731 (s, C=O), 1696 (s, C=O). ^{1}H NMR (Acetone-d_6), δ: 9.94 (s, NH), 7.81 (s, 1H, isatin_4 H), 7.66 (s, 1H, phenyl_2 H), 7.60 (s, 1H, isatin_6 H), 7.50 (d, *J* = 7.6 Hz, 1H, pheny_6 H), 7.40 (t, *J* = 7.6 Hz, 1H, phenyl_5 H), 7.29 (d, *J* = 7.6 Hz, 1H, phenyl_4 H), 3.03 (t, *J* = 7.3 Hz, 2H, COOH-C$\underline{H}_2$-CH_2), 2.71 (t, *J* = 7.3 Hz, 2H, COOH-CH_2-C$\underline{H}_2$), 2.41 (s, 3H, CH_3). ^{13}C NMR (DMSO-d_6), δ: 184.8 ($\underline{C}$OOH), 173.9 (isatin C3), 160.1 (isatin C2), 148.5 (isatin C7a), 141.7 (phenyl C3), 138.7 (phenyl C1), 137.6 (isatin C6), 134.8 (isatin C5), 128.9 (phenyl C5), 127.4 (phenyl C6), 126.5 (phenyl C2), 123.8 (phenyl C4), 122.1 (isatin C4), 119.8 (isatin C3a), 118.1 (isatin C7), 35.3 (COOH-$\underline{C}H_2$-CH_2), 30.4 (COOH-CH_2-$\underline{C}H_2$), 15.5 (CH_3) ppm. EI-MS *m/z* 309 (M+, 100%), 277 (25), 253 (80), 207 (79), 193 (30). HRMS (ESI) calcd for $C_{18}H_{16}NO_4$ (MH+), 310.1079; found, 310.1071. HPLC purity: 98.1%, R_t=5.339 (1% solvent B in solvent A).

3-(3-(4-Methoxy-7-methyl-2,3-dioxoindolin-5-yl)phenyl) propanoic acid ***142***



Following the general procedure 6 using 5-bromo-4-methoxy-7-methylindoline-2,3-dione **112** (100 mg, 0.370 mmol), 3-(3-boronophenyl)propanoic acid **113** (86.1 mg, 0.444 mmol), tetrakis(triphenylphosphine)palladium(0) (21.4 mg, 0.019 mmol) and potassium phosphate (235.6 mg, 1.11 mmol) and heating for 5 h, the acid **142** was isolated as a red solid (60 mg, 48%), mp: 219-220 ºC; IR (cm^{-1}): υ 3182 (m, NH), 2982 (m, OH), 1749 (s, C=O), 1692 (s, C=O), 1287 (w, C-O). ^{1}H NMR (DMSO-d_6), δ: 11.16 (s, NH), 7.40 (s, 1H, isatin_6 H), 7.13 (t, *J* = 7.4 Hz, 1H, phenyl_5 H), 7.26 (s, 1H, phenyl_2 H), 7.22 (d, *J* = 7.4 Hz, 1H, pheny_6 H), 7.19 (d, *J* = 7.4 Hz, 1H, phenyl_4 H), 3.81 (s, 3H, OCH_3), 2.86 (t, *J* = 7.4 Hz, 2H, COOH-C$\underline{H}_2$-CH_2), 2.56 (t, *J* = 7.4 Hz, 2H, COOH-CH_2-C$\underline{H}_2$), 2.41 (s, 3H, CH_3). ^{13}C NMR, δ: 182.8 (C$\underline{O}$OH), 174.5 (isatin C2), 160.5 (isatin C3), 154.6 (isatin C4), 149.3 (isatin C3a), 142.6 (phenyl C1), 141.5 (isatin C6), 137.4 (phenyl C3), 129.5 (phenyl C2), 129.1 (phenyl C5), 128.8 (isatin C7), 127.7 (phenyl C6), 127.3 (phenyl C4), 116.5 (isatin C5), 110.3 (isatin C3a), 62 (OCH_3), 35.9 (COOH-C$\underline{H}_2$-CH_2), 31.0 (COOH-CH_2-C$\underline{H}_2$), 15.6 (CH_3) ppm. EI-MS *m/z* 339 (M+, 100%), 311 (75), 293 (10), 277 (77). HRMS (ESI) calcd for $C_{19}H_{18}NO_5$ (MH+), 340.1185; found, 340.1187. HPLC purity: 99.4%, R_t=5.579 (1% solvent B in solvent A).

3-(3-(2,3-Dioxoindolin-5-yl)phenyl)propanoic acid ***143***



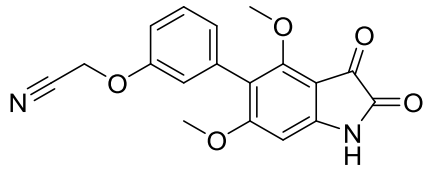
Following the general procedure 6 using 5-bromoindoline-2,3-dione **139** (100mg, 0.442 mmol), 3-(3-boronophenyl)propanoic acid **113** (102.9 mg, 0.530 mmol), tetrakis(triphenylphosphine)palladium(0) (25.5 mg, 0.022 mmol) and potassium

phosphate (281.5 mg, 1.326 mmol) and heating for 4 h, the acid **143** was isolated as an orange solid (84.8 mg, 65%), mp: 182-183 ºC; IR (cm^{-1}): υ 3185 (m, NH), 2989 (m, OH), 1714 (s, C=O), 1684 (s, C=O). ^{1}H NMR (DMSO-d_6), δ: 11.14 (s, NH), 7.89 (dd, *J* = 8.2, 1.8 Hz, 1H, isatin_6 H), 7.77 (d, *J* = 1.4 Hz, 1H, isatin_4 H), 7.53 (s, 1H, phenyl_2 H), 7.46 (d, *J* = 7.6 Hz, 1H, phenyl_4 H), 7.35 (t, *J* = 7.5 Hz, 1H, phenyl_5 H), 7.21 (d, *J* = 7.5 Hz, 1H, phenyl_6 H), 6.99 (d, *J* = 7.9 Hz, 1H, isatin_7 H), 2.88 (t, *J* = 7.6 Hz, 2H, COOH-C$\underline{H}_2$-CH$_2$), 2.59 (t, *J* = 7.5 Hz, 2H, COOH-CH$_2$-C$\underline{H}_2$). ^{13}C NMR, δ: 184.9 ($\underline{C}$OOH), 174.4 (isatin C3), 160.0 (isatin C2), 150.4 (isatin C7a), 142.2 (phenyl C3), 139.1 (phenyl C1), 136.9 (isatin C6), 135.4 (isatin C 5), 129.4 (phenyl C5), 127.9 (phenyl C6), 126.7 (phenyl C2), 124.4 (phenyl C4), 122.9 (isatin C4), 118.9 (isatin C3a), 113.1 (isatin C7), 35.7 (COOH-$\underline{C}$H$_2$-CH$_2$), 30.9 (COOH-CH$_2$-$\underline{C}$H$_2$) ppm. EI-MS *m/z* 295 (M+, 40), 267 (100%), 250 (95), 207 (20), 193 (25). HRMS (ESI) calcd for $C_{17}H_{14}NO_4$ (MH+), 296.0914; found, 296.0912. HPLC purity: 98.2%, R_t=7.237 (1% solvent B in solvent A).

General procedure 7: Suzuki coupling of the 5-bromoisatin derivatives 103, **106**, **112 and 139 with (3-(cyanomethoxy)phenyl)boronic acid 118:**

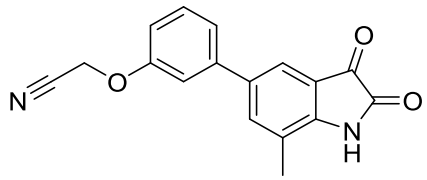
Under a N_2 atmosphere, a 25 mL flask was charged with the 5-bromoisatin derivative (1 equiv), (3-(cyanomethoxy)phenyl)boronic acid **118** (1 equiv), tetrakis(triphenylphosphine)palladium(0) (5 mol%), potassium phosphate (3 equiv) and a mixture of diglyme (3 mL) and water (3 mL), and the reaction mixture was degassed by flushing N_2 through a needle for 10 min. The flask was then sealed and heated at 85 ºC for 5-9 h. After cooling to room temperature, the solvent was concentrated under reduced pressure, and 1 N HCl (15 mL) was added. The suspended solid was either collected by filtration or extracted with ethyl acetate (2 x 25 mL). The combined organic layers were dried ($MgSO_4$) and evaporated under reduced pressure. The combined residues were adsorbed onto silica and subjected to flash column chromatography (40% ethyl acetate in petroleum spirit) to yield the acetonitrile derivative **144**-**147**.

2-(3-(4,6-Dimethoxy-2,3-dioxoindolin-5-yl)phenoxy)acetonitrile ***144***



Following the general procedure 7 using 5-bromo-4,6-dimethoxyindoline-2,3-dione **103** (40 mg, 0.139 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (24.6 mg, 0.139 mmol), tetrakis(triphenylphosphine)palladium(0) (8 mg, 0.007 mmol), potassium phosphate (88.5 mg, 0.417 mmol) and heating for 9 h, the acetonitrile intermediate **144** could not be isolated. Analysis of the crude mixture: EI-MS *m/z* 338 (M+, 17), 309 (7), 382 (66), 257 (100%), 225 (25). HRMS (ESI) calcd for $C_{18}H_{15}N_2O_5$ (MH+), 339.0981; found, 339.0980. This sample was collected from the reaction mixture before purification trial. Could not purify and was not stable enough to get NMR data.

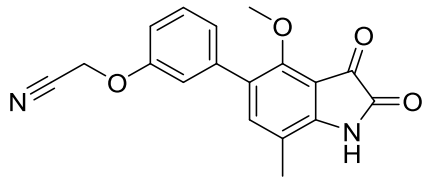
2-(3-(7-Methyl-2,3-dioxoindolin-5-yl)phenoxy)acetonitrile ***145***



Following the general procedure 7 using 5-bromo-7-methylindoline-2,3-dione **106** (100 mg, 0.416 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (73.6 mg, 0.416 mmol), tetrakis(triphenylphosphine)palladium(0) (24 mg, 0.021 mmol), potassium phosphate (264.9 mg, 1.248 mmol) and heating for 5 h, the acetonitrile intermediate **145** was isolated as a red solid (63.2 mg, 52%), mp: 208-209 ºC; IR (cm^{-1}): υ 3185 (w, NH), 2365 (m, CN), 1750 (s, C=O), 1587 (s, C=O). ^{1}H NMR (Acetone-d_6), δ: 10.13 (s, NH), 7.83 (s, 1H, isatin_4 H), 7.68 (s, 1H, isatin_C6), 7.47 (t, *J* = 7.8 Hz, 1H, phenoxy_5 H), 7.40-7.38 (m, 2H, phenoxy_2,4 H), 7.09 (d, *J* = 8.0 Hz, 1H, phenoxy_6 H), 5.21 (s, 2H, CH_2), 2.39 (s, 3H, CH_3). ^{13}C NMR, δ: 185.1 (isatin C3), 160.5 (phenoxy C1), 158.6 (isatin C2), 149.8 (isatin C7a), 142.4 (phenoxy C3), 138.9 (isatin C7), 136.4 (isatin C6), 131.3 (isatin C5), 123.3 (phenoxy C5), 121.8 (isatin C4), 121.2 (phenoxy C4), 119.5 (phenoxy C2), 116.7 (phenoxy C6), 115.3 (CN), 114.1 (isatin C3a), 54.7 (CH_2), 15.8

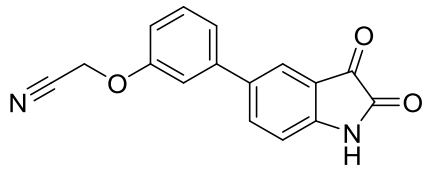
(CH_3) ppm. EI-MS *m/z* 292 (M+, 93), 264 (80), 236 (100%), 196 (17), 180 (11). HRMS (ESI) calcd for $C_{17}H_{13}N_2O_3$ (MH+), 293.0926; found, 293.0931. HPLC purity: 97.9%, R_t=13.070 (2.5% solvent C in solvent A).

2-(3-(4-Methoxy-7-methyl-2,3-dioxoindolin-5-yl)phenoxy) acetonitrile ***146***



Following the general procedure 7 using 5-bromo-4-methoxy-7-methylindoline-2,3-dione **112** (100 mg, 0.370 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (65.5 mg, 0.370 mmol), tetrakis(triphenylphosphine)palladium(0) (21.4 mg, 0.019 mmol), potassium phosphate (235.6 mg, 1.11 mmol) and heating for 8 h, the acetonitrile intermediate **146** was isolated as a red solid (41.7 mg, 35%), mp: 179-180 ºC; IR (cm^{-1}): υ 3190 (m, NH), 2166 (w, CN), 1706 (s, C=O), 1611 (s, C=O), 1294 (w, C-O). ^{1}H NMR (CD_3OD), δ: 7.40 (s, 1H, isatin_6 H), 7.38 (t, *J* = 8.2 Hz, 1H, phenoxy_5 H), 7.12 (d, *J* = 8.2 Hz, 1H, phenoxy_4 H), 7.11 (s, 1H, phenoxy_2 H), 7.01 (d, *J* = 8.3 Hz, 1H, phenoxy_6 H), 5.00 (s, 2H, CH_2), 3.91 (s, 3H, OCH_3), 2.22 (s, 3H, CH_3). ^{13}C NMR, δ: 160.2 (isatin C3), 156.7 (phenoxy C1), 154.9 (isatin C2), 148.6 (isatin C4), 142.1 (isatin C7a), 138.9 (isatin C6), 137.8 (phenoxy C3), 129.2 (phenoxy C5), 123.4 (isatin C5), 120.5 (isatin C7), 115.9 (phenoxy C4), 115.7 (phenoxy C6), 115.5 (phenoxy C2), 113.7 (CN), 109.5 (isatin C3a), 61.13 (CH_2), 53.3 (OCH_3), 13.7 (CH_3) ppm. EI-MS *m/z* 322 (M+, 74), 294 (68), 280 (4), 254 (7), 226 (100%), 211 (11). HRMS (ESI) calcd for $C_{18}H_{15}N_2O_4$ (MH+), 323.1032; found, 323.1018. HPLC purity: 99.5%, R_t=11.766 (2.5% isopropanol in solvent A).

*2-(3-(2,3-Dioxoindolin-5-yl)phenoxy)acetonitrile **147***



Following the general procedure 7 using 5-bromoindoline-2,3-dione **139** (100 mg, 0.442 mmol), (3-(cyanomethoxy)phenyl)boronic acid **118** (78.2 mg, 0.442 mmol), tetrakis(triphenylphosphine)palladium(0) (25.5 mg, 0.022 mmol), potassium phosphate (281.5 mg, 1.326 mmol) and heating for 5 h, the acetonitrile intermediate **147** was isolated as a red solid (73.7 mg, 60%), mp: 205-206 ºC; IR (cm^{-1}): υ 3168 (w, NH), 2166 (m, CN), 1732 (s, C=O), 1621 (s, C=O). ^{1}H NMR (Acetone-d_6), δ: 10.04 (s, NH), 7.95 (d, *J* = 8.2, 1H, phenoxy_4 H), 7.83 (s, 1H, phenoxy_2 H), 7.47 (t, *J* = 8.2 Hz, 1H, phenoxy_5 H), 7.39-7.38 (m, 2H, isatin_4,7 H), 7.14 (d, *J* = 8.2 Hz, phenoxy_6 H), 7.09 (d, *J* = 8.6 Hz, 1H, isatin_6 H), 5.21 (s, 2H, CH_2). ^{13}C NMR, δ: 184.1 (isatin C3), 159.6 (phenoxy C1), 157.8 (isatin C2), 150.4 (isatin C7a), 141.5 (phenoxy C3), 137.0 (phenoxy C4), 135.7 (phenoxy C2), 130.6 (phenoxy C5), 123.1 (isatin C5), 121.0 (isatin C7), 118.9 (phenoxy C6), 116.0 (CN), 114.6 (isatin C3a), 113.1 (isatin C4), 112.9 (isatin C6), 53.8 (CH_2) ppm. EI-MS *m/z* 278 (M+, 46), 250 (100%), 207 (5), 182 (49). HRMS (ESI) calcd for $C_{16}H_{11}N_2O_3$ (MH+), 279.0770; found, 279.0759. HPLC purity: 99.7%, R_t=20.046 (2.5% solvent C in solvent A).

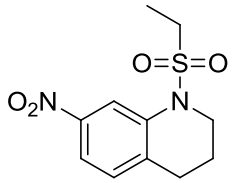
9.1.5. Experimental for chapter 7

General procedure 8: synthesis of compounds 162, 164-166

A solution of 7-nitro-1,2,3,4-tetrahydroquinoline **161** (1.0 g, 5.62 mmol) in dry pyridine (3 mL) was cooled to 0 ºC. The appropriate sulfonyl chloride (6.18 mmol) was then added dropwise over 10 min with stirring and the mixture was then allowed to warm to room temperature and stirred for 24 h. Cold water (20 mL) was added, the resulting solid or oil was extracted with ethyl acetate (2 x 25 mL). The combined organic layers were washed with 1 M HCl (2 x 20 mL), brine (10 mL) and dried ($MgSO_4$). The solvent was then removed under reduced pressure. The resulting residue was subjected to flash

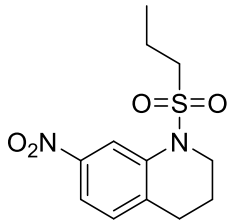
column chromatography and elution with 20% ethyl acetate in petroleum spirit yielded the corresponding intermediate.

*1-(Ethylsulfonyl)-7-nitro-1,2,3,4-tetrahydroquinoline **162***



Following general procedure 8 using ethanesulfonyl chloride (794 mg, 6.18 mmol) gave the intermediate **162** (1.1 g, 70%) as a yellow solid, mp: 64-65 ºC; IR (cm^{-1}): υ 3139 (w, C-H aromatic), 2944 (w, C-H aliphatic), 1517 (s, NO_2), 1345 (m, SO_2), 1333 (m, NO_2), 1147 (w, C-N). ^{1}H NMR, δ: 8.49 (s, 1H, H8), 7.86 (d, *J* = 8.3 Hz, 1H, H6), 7.28 (d, *J* = 8.3 Hz, 1H, H5), 3.83 (t, *J* = 6.0 Hz, 2H, H2), 3.22 (q, *J* = 7.5 Hz, 2H, ethyl CH_2), 2.95 (t, *J* = 6.5 Hz, 2H, H4), 2.07 (p, *J* = 6.0 Hz, 2H, H3), 1.41 (t, *J* = 7.5 Hz, 3H, CH_3). ^{13}C NMR, δ: 146.7 (C8a), 137.9 (C7), 135.7 (C5), 130.4 (C4a), 118.3 (C6), 116.8 (C8), 47.1 (ethyl CH_2), 46.4 (C2), 27.6 (C4), 22.1 (C3), 8.0 (CH_3) ppm. EI-MS *m/z* 270 (M+, 55), 177 (90), 131 (100%). HRMS (ESI) calcd for $C_{11}H_{15}N_2O_4S$ (MH+), 271.0753; found, 271.0744.

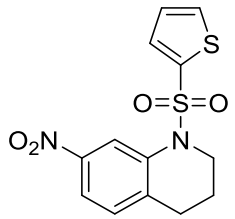
*7-Nitro-1-(propylsulfonyl)-1,2,3,4-tetrahydroquinoline **164***



Following general procedure 8 using propane-1-sulfonyl chloride (880.5 mg, 6.18 mmol) gave the intermediate **164** (1.28 g, 80%) as a yellow solid, mp: 79-80 ºC; IR (cm^{-1}): υ 3128 (w, C-H aromatic), 2972 (w, C-H aliphatic), 1517 (s, NO_2), 1346 (m, SO_2), 1332 (m, NO_2), 1143 (w, C-N). ^{1}H NMR, δ: 8.51 (s, 1H, H8), 7.86 (d, *J* = 8.3 Hz, 1H, H6), 7.28 (d, *J* = 8.4 Hz, 1H, H5), 3.83 (t, *J* = 5.8 Hz, 2H, H2), 3.13 (t, *J* = 7.7 Hz, 2H,

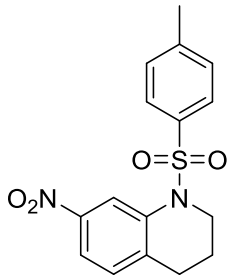
propyl_1 CH_2), 2.95 (t, *J* = 6.4 Hz, 2H, H4), 2.07 (p, *J* = 5.9 Hz, 2H, H3), 1.89 (sex, *J* = 7.6 Hz, 2H, propyl_2 CH_2), 1.06 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 146.7 (C8a), 137.9 (C7), 135.8 (C5), 130.4 (C4a), 118.3 (C6), 116.8 (C8), 54.3 (propyl C1), 46.3 (C2), 27.6 (C4), 22.1 (C3), 17.1 (propyl C2), 12.9 (CH_3) ppm. EI-MS *m/z* 284 (M+, 54), 178 (92), 131 (100%). HRMS (ESI) calcd for $C_{12}H_{17}N_2O_4S$ (MH+), 285.0909; found, 285.0916.

7-Nitro-1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinoline ***165***



Following general procedure 8 using thiophene-2-sulfonyl chloride (1.13 g, 6.18 mmol) gave the intermediate **165** (1.52 g, 84%) as an orange solid, mp: 130-131 ºC; IR (cm^{-1}): υ 3139 (w, C-H aromatic), 2966 (w, C-H aliphatic), 1515 (s, NO_2), 1341 (m, SO_2), 1292 (m, NO_2), 1160 (w, C-N). ^{1}H NMR, δ: 8.72 (s, 1H, H8), 7.91 (d, *J* = 8.4 Hz, 1H, H6), 7.56 (d, *J* = 4.8 Hz, 1H, thiophene_5 H), 7.52 (d, *J* = 3.1 Hz, 1H, thiophene_3 H), 7.21 (d, *J* = 8.4 Hz, 1H, H5), 7.06 (t, *J* = 4.4 Hz, 1H, thiophene_4 H), 3.89 (t, *J* = 6.2 Hz, 2H, H2), 2.66 (t, *J* = 6.5 Hz, 2H, H4), 1.80 (p, *J* = 6.2 Hz, 2H, H3). ^{13}C NMR, δ: 146.7 (C8a), 138.9 (thiophene C1), 137.5 (C7), 137.2 (C5), 133.0 (thiophene C3), 132.9 (C4a), 129.9 (thiophene C4), 127.6 (thiophene C5), 119.4 (C6), 119.3 (C8), 46.6 (C2), 27.3 (C4), 21.2 (C3) ppm. EI-MS *m/z* 324 (M+, 10), 243 (9), 214 (17), 177 (15), 131 (100%). HRMS (ESI) calcd for $C_{13}H_{13}N_2O_4S_2$ (MH+), 325.0317; found, 325.0325.

*7-Nitro-1-tosyl-1,2,3,4-tetrahydroquinoline **166***

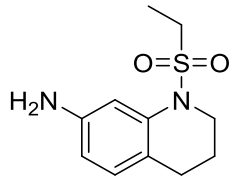


Following general procedure 8 using 4-methylbenzenesulfonyl chloride (1.18 g, 6.18 mmol) gave the intermediate **166** (1.64 g, 88%) as a yellow crystalline solid, mp: 133-134 ºC; IR (cm^{-1}): υ 3136 (m, C-H aromatic), 2949 (m, C-H aliphatic), 1510 (s, NO_2), 1340 (m, SO_2), 1334 (w, NO_2), 1163 (w, C-N). ^{1}H NMR, δ: 8.68 (s, 1H, H8), 7.87 (dd, *J* = 8.4, 2.1 Hz, 1H, H6), 7.57 (d, *J* = 8.3 Hz, 2H, tosyl_2,6 H), 7.25 (d, *J* = 8.1 Hz, 2 H, tosyl_3,5 H), 7.17 (d, *J* = 8.4 Hz, 1H, H5), 3.85 (t, *J* = 6.0 Hz, 2H, H2), 2.64 (t, *J* = 6.5 Hz, 2H, H4), 2.40 (s, 3H, CH_3), 1.75 (p, *J* = 6.5 Hz, 2H, H3). ^{13}C NMR, δ: 146.6 (C8a), 144.4 (C7), 137.8 (tosyl C4), 137.0 (tosyl C1), 136.1 (C5), 129.93 (C4a), 129.90 (tosyl C3, C5), 127.2 (tosyl C2, C6), 118.9 (C6), 118.8 (C8), 46.3 (C2), 27.2 (C4), 21.6 (C3), 21.1 (CH_3) ppm. EI-MS *m/z* 332 (M+, 70), 268 (46), 251 (49), 177 (70), 130 (100%). HRMS (ESI) calcd for $C_{16}H_{17}N_2O_4S$ (MH+), 333.0909; found, 333.0900.

General procedure 9: synthesis of compounds 163, 167-169

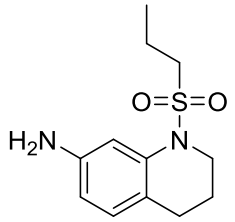
A solution of the appropriate sulfonamide **162**, **164-166** (1 equiv) and Raney Ni (0.05 equiv) in methanol (20 mL) was stirred at room temperature under a N_2 atmosphere. A solution of hydrazine hydrate (1.2 equiv) in methanol (10 mL) was then added dropwise and stirring continued for 10 min, after which the H_2 gas ceased to evolve, and the reaction mixture was then heated at reflux for 4-6 h. The mixture was then filtered hot through a short pad of celite, and the filterate was concentrated under reduced pressure. The resulting residues were recrystallised either from methanol or toluene (80-90%).

*7-Amino-1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinoline **163***



Following the general procedure 9 using 1-(ethylsulfonyl)-7-nitro-1,2,3,4-tetrahydroquinoline **162** (1.1 g, 4.07 mmol), hydrazine hydrate (244 mg, 4.884 mmol) and heating for 4 h, the amine intermediate **163** was obtained by recrystallisation from methanol as a white solid (763 mg, 78%), mp: 65-66 ºC; IR (cm^{-1}): υ 3358 (w, C-H aromatic), 2943 (w, C-H aliphatic), 1323 (m, SO_2), 1143 (w, C-N). ^{1}H NMR, δ: 7.07 (s, 1H, H8), 6.88 (d, *J* = 8.0 Hz, 1H, H6), 6.40 (d, *J* = 8.0 Hz, 1H, H5), 3.78 (t, *J* = 6.0 Hz, 2H, H2), 3.62 (s, br, NH_2), 3.10 (q, *J* = 7.3 Hz, 2H, ethyl CH_2), 2.73 (t, *J* = 6.4 Hz, 2H, H4), 1.95 (p, *J* = 6.0 Hz, 2H, H3), 1.33 (t, *J* = 7.3 Hz, 3H, CH_3). ^{13}C NMR, δ: 145.4 (C7), 137.9 (C8a), 130.6 (C5), 118.7 (C4a), 111.7 (C6), 108.3 (C8), 46.9 (ethyl C1), 46.2 (C2), 26.7 (C4), 23.1 (C3), 8.2 (CH_3) ppm. EI-MS *m/z* 240 (M+, 30), 147 (100%), 130 (49). HRMS (ESI) calcd for $C_{11}H_{17}N_2O_2S$ (MH+), 241.1011; found, 241.1005.

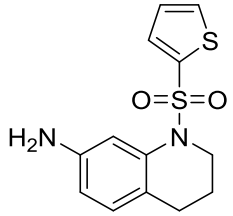
*7-Amino-1-(propylsulfonyl)-1,2,3,4-tetrahydroquinoline **167***



Following the general procedure 9 using 7-nitro-1-(propylsulfonyl)-1,2,3,4-tetrahydroquinoline **164** (1.2 g, 4.22 mmol), hydrazine hydrate (253 mg, 5.064 mmol) and heating for 4 h, the amine intermediate **167** was obtained by recrystallisation from methanol as an orange solid (808 mg, 75%), mp: 86-87 ºC; IR (cm^{-1}): υ 3358 (m, NH_2), 2941 (w, C-H aliphatic), 1327 (m, SO_2), 1139 (m, C-N). ^{1}H NMR, δ: 7.07 (s, 1H, H8), 6.88 (d, *J* = 8.1 Hz, 1H, H6), 6.40 (d, *J* = 7.8 Hz, 1H, H5), 3.77 (t, *J* = 5.8 Hz, 2H, H2), 3.65 (s, br, NH_2), 3.02 (t, *J* = 7.4 Hz, 2H, propyl_1 CH_2), 2.71 (t, *J* = 6.3 Hz, 2H, H4), 1.93 (p, *J* = 6.0 Hz, 2H, H3), 1.83 (sex, *J* = 7.5 Hz, 2H, propyl_2 CH_2), 1.00 (t, *J* = 7.4

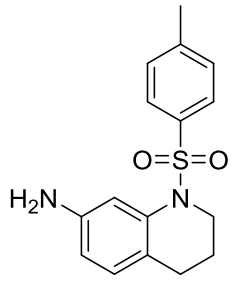
Hz, 3H, CH_3). ^{13}C NMR, δ: 145.2 (C8), 137.6 (C8a), 130.4 (C5), 118.4 (C4a), 111.5 (C6), 108.2 (C8), 53.3 (propyl C1), 46.6 (C1), 26.5 (C4), 22.8 (C3), 17.0 (propyl C2), 13.0 (CH_3) ppm. ESI-MS *m/z* 255.2 (M+H). HRMS (ESI) calcd for $C_{12}H_{19}N_2O_2S$ (MH+), 255.1167; found, 255.1156.

7-Amino-1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinoline ***168***



Following the general procedure 9 using 7-nitro-1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinoline **165** (1.0 g, 3.09 mmol), hydrazine hydrate (185 mg, 3.708 mmol) and heating for 6 h, the amine intermediate **168** was obtained by recrystallisation from toluene as a yellow solid (762 mg, 84%), mp: 78-79 ºC; IR (cm^{-1}): υ 3363 (w, NH_2), 3092 (w, C-H aromatic), 2937 (w, C-H aliphatic), 1337 (m, SO_2), 1153 (w, C-N). ^1H NMR, δ: 7.49 (d, *J* = 4.4 Hz, 1H, thiophene_5 H), 7.38 (d, *J* = 3.8 Hz, 1H, thiophene_3 H), 7.19 (s, 1H, H8), 6.99 (t, *J* = 4.5 Hz, 1H, thiophene_4 H), 6.80 (d, *J* = 8.0 Hz, 1H, H5), 6.46 (dd, *J* = 8.0, 2.1 Hz, 1H, H6), 3.80 (t, *J* = 6.0 Hz, 2H, H2), 3.67 (s, br, NH_2) 2.38 (t, *J* = 6.5 Hz, 2H, H4), 1.63 (p, *J* = 6.3 Hz, 2H, H3). ^{13}C NMR, δ: 144.9 (C7), 139.6 (C8a), 136.8 (thiophene C1), 131.9 (thiophene C3), 131.8 (C4a), 129.8 (C5), 127.2 (thiophene C4), 120.7 (thiophene C5), 112.8 (C6), 111.4 (C8), 47.1 (C2), 25.9 (C4), 21.5 (C3) ppm. EI-MS *m/z* 294 (M+, 47), 230 (100%), 145 (52), 130 (65). HRMS (ESI) calcd for $C_{13}H_{15}N_2O_2S_2$ (MH+), 295.0575; found, 295.0562.

*7-Amino-1-tosyl-1,2,3,4-tetrahydroquinoline **169***



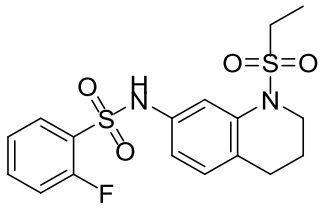
Following the general procedure 9 using 7-nitro-1-tosyl-1,2,3,4-tetrahydroquinoline **166** (1.0 g, 3.01 mmol), hydrazine hydrate (180.6 mg, 3.612 mmol) and heating for 6 h, the amine intermediate **169** was obtained by recrystallisation from toluene as a yellow crystalline solid (773 mg, 85%), mp: 89-90 ºC (Lit.[266] m.p: 109-110 ºC); IR (cm^{-1}): υ 3353 (w, NH_2), 2949 (w, C-H aliphatic), 1339 (m, SO_2), 1156 (w, C-N). ^{1}H NMR, δ: 7.50 (d, *J* = 8.1 Hz, 2H, tosyl_3,5 H), 7.20 (s, 1H, H8), 7.18 (d, *J* = 8.1 Hz, 2H, tosyl_2,6 H), 6.76 (d, *J* = 8.1 Hz, 1H, H5), 6.44 (dd, *J* = 8.0, 2.1 Hz, 1H, H6), 3.76 (t, *J* = 5.9 Hz, 2H, H2), 3.64 (s, br, NH_2), 2.37 (s, 3H, CH_3), 2.33 (t, *J* = 6.4 Hz, 2H, H4), 1.55 (p, *J* = 6.1 Hz 2H, H3). ^{13}C NMR, δ: 144.9 (C7), 143.4 (C8a), 137.4 (tosyl C4), 136.9 (tosyl C1), 129.5 (C5), 127.13 (C4a), 127.07 (tosyl C3, C5), 120.4 (tosyl C2, C6), 112.5 (C6), 111.2 (C8), 46.7 (C2), 25.8 (C4), 21.59 (C3), 21.51 (CH_3) ppm. EI-MS *m/z* 302 (M+, 55), 238 (46), 147 (100%), 130 (75). HRMS (ESI) calcd for $C_{16}H_{19}N_2O_2S$ (MH+), 303.1167; found, 303.1156.

General procedure 10: synthesis of compounds 160, **170-194**

A solution of the appropriate amine intermediate **163**, **167-169** (1 mmol) in dry pyridine (2 mL) was cooled at 0 ºC. The appropriate sulfonyl chloride (1.2 mmol) was added portion wise over a 10 min period with stirring. The reaction mixture was then allowed to warm to room temperature and stirred for 6 h. Cold water (5 mL) was then added and the resulting solid or oil was extracted with ethyl acetate (2 x 10 mL). The combined organic layers were washed with 1 M HCl (2 x 10 mL), brine (10 mL) and dried ($MgSO_4$). The solvent was then concentrated under reduced pressure and the resulting residues were subjected to flash column chromatography (30% ethylacetate in

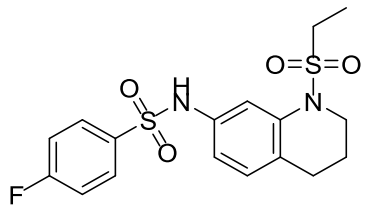
petroleum spirit) to yield the corresponding bis-sulfonamides **160** and **170**-**194** (70-80%).

*N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2-fluorobenzenesulfonamide **160***



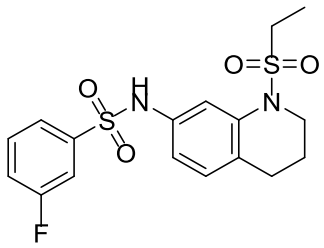
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 2-fluorobenzenesulfonyl chloride (233.5 mg), **160** was isolated as a colourless solid (282.6 mg, 71%), mp: 108-109 ºC; IR (cm^{-1}): υ 3223 (m, NH), 2949 (w, C-H aliphatic), 1472 (w, C-F), 1329 (m, SO_2), 1135 (w, C-N). ^{1}H NMR, δ: 7.86 (t, *J* = 6.9 Hz, 1H, benzene_4 H), 7.53 (q, *J* = 6.8 Hz, 1H, benzene_3 H), 7.42 (s, br, NH), 7.26-7.17 (m, 2H, benzene_5,6 H), 6.96 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 6.92 (s, 1H, tetrahydroquinoline_8 H), 6.86 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 3.73 (t, *J* = 6.0 Hz, 2H, tetrahydroquinoline_2 CH_2), 3.03 (q, *J* = 7.5 Hz, 2H, ethyl CH_2), 2.73 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 CH_2), 1.92 (p, *J* = 6.3 Hz, 2H, tetrahydroquinoline_3 CH_2), 1.27 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 159.6 (d, *J* = 254.9 Hz, benzene C2), 137.7 (tetrahydroquinoline C8a), 135.4 (d, *J* = 8.4 Hz, benzene C3), 134.4 (tetrahydroquinoline C7), 131.1 (tetrahydroquinoline C5), 130.1 (benzene C5), 126.4 (d, *J* = 13.0 Hz, benzene C1), 125.6 (benzene C6), 124.5 (d, *J* = 3.7 Hz, benzene C4), 117.1 (tetrahydroquinoline C4a), 116.7 (tetrahydroquinoline C6), 114.2 (tetrahydroquinoline C8), 46.5 (tetrahydroquinoline C2), 46.4 (ethyl CH_2), 26.7 (tetrahydroquinoline C4), 22.4 (tetrahydroquinoline C3), 8.0 (CH_3) ppm. EI-MS *m/z* 398 (M+, 50), 305 (45), 159 (15), 145 (100%). HRMS (ESI) calcd for $C_{17}H_{20}FN_2O_4S_2$ (MH+), 399.0849; found, 399.0862. HPLC purity: 99.4%, R_t=12.672 (20% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-4-fluorobenzenesulfonamide ***170***



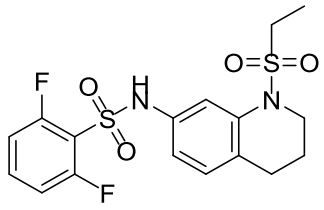
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg 1.0 mmol), 4-fluorobenzenesulfonyl chloride (233.5 mg), **170** was isolated as a colourless solid (286.6 mg, 72%), mp: 90-91 ºC; IR (cm^{-1}): υ 3276 (m, NH), 2955 (w, C-H aliphatic), 1456 (w, C-F), 1327 (m, SO_2), 1166 (w, C-N). ^{1}H NMR, δ: 7.84 (dd, *J* = 8.2, 5.2 Hz, 2H, benzene_3,5 H), 7.35 (s, 1H, tetrahydroquinoline_8 H), 7.12 (s, br, NH), 7.09 (d, *J* = 8.2 Hz, 2H, benzene_2,6 H), 6.99 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.86 (d, *J* = 8.1 Hz, 1H, tetrahydroquinoline_6 H), 3.75 (t, *J* = 6.0 Hz, 2H, tetrahydroquinoline_2 H), 3.05 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.76 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.96 (p, *J* = 6.0 Hz, 2H, tetrahydroquinoline_3 H), 1.27 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 165.5 (d, *J* = 266.4 Hz, benzene C4), 138.1 (tetrahydroquinoline C8a), 135.4 (d, *J* = 13.8 Hz, benzene C2, C6), 135.2 (tetrahydroquinoline C7), 130.9 (tetrahydroquinoline C4a), 130.6 (d, *J* = 10.1 Hz, benzene C1), 125.8 (tetrahydroquinoline C5), 117.0 (d, *J* = 90.3 Hz, benzene C3, C5), 116.5 (tetrahydroquinoline C8), 114.6 (tetrahydroquinoline C6), 46.9 (tetrahydroquinoline C2), 46.8 (ethyl CH_2), 27.0 (tetrahydroquinoline C4), 22.8 (tetrahydroquinoline C3), 8.3 (CH_3) ppm. EI-MS *m/z* 398 (M+, 17), 305 (10), 159 (20), 145 (65), 95 (100%). HRMS (ESI) calcd for $C_{17}H_{20}FN_2O_4S_2$ (MH+), 399.0849; found, 399.0846. HPLC purity: 97.3%, R_t=4.416 (20% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-3-fluorobenzenesulfonamide ***171***



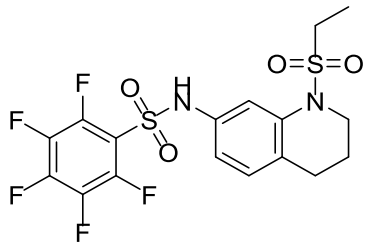
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 3-fluorobenzenesulfonyl chloride (233.5 mg), **171** was isolated as a pale yellow solid (274 mg, 69%), mp: 91-92 ºC; IR (cm^{-1}): υ 3272 (m, NH), 2955 (m, C-H aliphatic), 1451 (w, C-F), 1327 (m, SO_2), 1144 (w, C-N). ^{1}H NMR, δ: 7.64 (d, *J* = 7.6 Hz, 1H, benzene_2 H), 7.50 (d, *J* = 7.6 Hz, 1H, benzene_6 H), 7.44-7.40 (m, 2H, benzene_4,5 H), 7.37 (s, 1H, tetrahydroquinoline_8 H), 7.21 (t, *J* = 8.1 Hz, 1H, benzene_5 H), 6.99 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.89 (d, *J* = 8.1 Hz, 1H, tetrahydroquinoline_6 H), 3.74 (t, *J* = 5.5 Hz, 2H, tetrahydroquinoline_2 H), 3.06 (q, *J* = 7.2 Hz, 2H, ethyl CH_2), 2.75 (t, *J* = 5.9 Hz, 2H, tetrahydroquinoline_4 H), 1.95 (p, *J* = 5.5 Hz, 2H, tetrahydroquinoline_3 H), 1.27 (t, *J* = 7.2 Hz, 3H, CH_3). ^{13}C NMR, δ: 162.5 (d, *J* = 251.2 Hz, benzene C3), 141.3 (d, *J* = 6.5 Hz, benzene C1), 137.8 (tetrahydroquinoline C8a), 135.1 (tetrahydroquinoline C7), 131.2 (d, *J* = 7.5 Hz, benzene C5), 130.8 (tetrahydroquinoline C5), 125.8 (tetrahydroquinoline C4a), 123.5 (d, *J* = 3.7 Hz, benzene C6), 120.4 (d, *J* = 21.4 Hz, benzene C4), 117.5 (tetrahydroquinoline C6), 115.0 (d, *J* = 25.1 Hz, benzene C2), 114.7 (tetrahydroquinoline C8), 46.8 (tetrahydroquinoline C2), 46.7 (ethyl CH_2), 26.9 (tetrahydroquinoline C4), 22.7 (tetrahydroquinoline C3), 8.2 (CH_3) ppm. EI-MS *m/z* 398 (M+, 33), 305 (33), 159 (10), 145 (80), 118 (60), 95 (100%). HRMS (ESI) calcd for $C_{17}H_{20}FN_2O_4S_2$ (MH+), 399.0849; found, 399.0837. HPLC purity: 99.3%, R_t=10.956 (20% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2,6-difluorobenzenesulfonamide ***172***



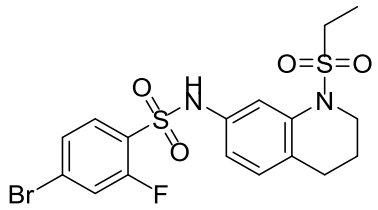
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 2,6-difluorobenzenesulfonyl chloride (255 mg), **172** was isolated as a white solid (295 mg, 71%), mp: 118-119 ºC; IR (cm^{-1}): υ 3245 (m, NH), 2942 (w, C-H aliphatic), 1467 (w, C-F), 1318 (m, SO_2), 1139 (w, C-N). ^1H NMR (CD_3OD), δ: 7.62-7.54 (m, 1H, benzene_4 H), 7.47 (s, 1H, tetrahydroquinoline_8 H), 7.07 (t, *J* = 9.0 Hz, 2H, benzene_3,5 H), 7.02 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.87 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 3.71 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline_2 H), 3.09 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.75 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 1.93 (p, *J* = 6.4 Hz, 2H, tetrahydroquinoline_3 H), 1.23 (t, *J* = 7.3 Hz, 3H, CH_3). ^{13}C NMR (CD_3OD), δ: 160.0 (dd, *J* = 256.5, 3.8 Hz, benzene C2, C6), 139.5 (tetrahydroquinoline C8a), 137.9 (tetrahydroquinoline C7), 135.2 (t, *J* = 10.9 Hz, benzene C1), 130.2 (benzene C4), 125.7 (tetrahydroquinoline C5), 117.1 (tetrahydroquinoline C4a), 116.5 (tetrahydroquinoline C6), 114.0 (tetrahydroquinoline C8), 113.0 (dd, *J* = 23.3, 3.4 Hz, benzene C3, C5), 46.4 (tetrahydroquinoline C2), 46.0 (ethyl CH_2), 26.4 (tetrahydroquinoline C4), 22.4 (tetrahydroquinoline C3), 7.0 (CH_3) ppm. EI-MS *m/z* 416 (M+, 30), 323 (27), 177 (20), 145 (100%), 118 (82). HRMS (ESI) calcd for $C_{17}H_{19}F_2N_2O_4S_2$ (MH+), 417.0754; found, 417.0742. HPLC purity: 99.3%, R_t=18.012 (15% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2,3,4,5,6-pentafluoro benzenesulfonamide ***173***



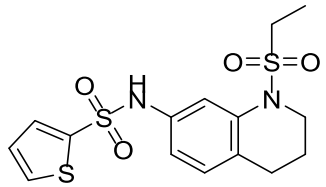
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 2,3,4,5,6-pentafluorobenzenesulfonyl chloride (319.8 mg), **173** was isolated as a white solid (348 mg, 74%), mp: 122-123 ºC; IR (cm^{-1}): υ 3201 (m, NH), 2947 (w, C-H aliphatic), 1464 (m, C-F), 1324 (m, SO_2), 1143 (w, C-N). ^{1}H NMR, δ: 7.39 (s, br, NH), 7.32 (s, 1H, tetrahydroquinoline_8 H), 7.06 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.92 (dd, *J* = 8.5, 1.9 Hz, 1H, tetrahydroquinoline_6 H), 3.74 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2), 3.12 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.79 (t, *J* = 6.6 Hz, 2H, tetrahydroquinoline_4 H), 1.97 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.34 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 145.1 (ddd, *J* = 259.2, 12.1, 4.7 Hz, benzene C4), 144.6 (dt, *J* = 263.3, 13.1 Hz, benzene C2, C6), 138.3 (dt, *J* = 258.6, 12.1 Hz, benzene C3, C5), 138.2 (tetrahydroquinoline C8a), 137.5 (d, *J* = 117.2 Hz, benzene C1), 133.4 (tetrahydroquinoline C7), 131.2 (tetrahydroquinoline C5), 127.2 (tetrahydroquinoline C4a), 117.8 (tetrahydroquinoline C6), 115.0 (tetrahydroquinoline C8), 47.2 (tetrahydroquinoline C2), 46.7 (ethyl CH_2), 27.0 (tetrahydroquinoline C4), 22.7 (tetrahydroquinoline C3), 8.2 (CH_3) ppm. ^{19}F NMR, δ: -136.6 (2F), -145.0, -158.6 (2F). ESI-MS *m/z* 469.2 (M-H). HRMS (ESI) calcd for $C_{17}H_{16}F_5N_2O_4S_2$ (MH+), 471.0472; found, 471.0452. HPLC purity: 95.3%, R_t=6.211 (20% solvent D in solvent A).

4-Bromo-N-(1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2-fluoro benzenesulfonamide ***174***



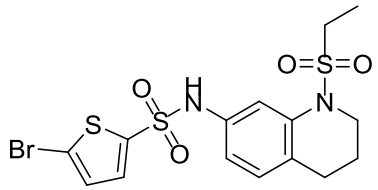
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 4-bromo-2-fluorobenzenesulfonyl chloride (328.2 mg), **174** was isolated as a red solid (328 mg, 69%), mp: 138-139 ºC; IR (cm^{-1}): υ 3244 (w, NH), 2967 (w, C-H aliphatic), 1471 (w, C-F), 1314 (m, SO_2), 1130 (w, C-N). ^{1}H NMR, δ: 7.73 (t, *J* = 7.9 Hz, 1H, benzene_3 H), 7.40-7.37 (m, 3H, benzene_5,6, tetrahydroquinoline_8 H), 6.99 (d, *J* = 8.1 Hz, 1H, tetrahydroquinoline_5 H), 6.87 (s, br, NH), 6.84 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 3.74 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.03 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.75 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.94 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.29 (t, *J* = 7.3 Hz, 3H, CH_3). ^{13}C NMR, δ: 157.5 (d, *J* = 259.6 Hz, benzene C2), 138.0 (tetrahydroquinoline C8a), 134.1 (tetrahydroquinoline C7), 132.3 (tetrahydroquinoline C5), 130.7 (tetrahydroquinoline C4a), 129.0 (d, *J* = 9.3 Hz, benzene C6), 128.2 (d, *J* = 3.7 Hz, benzene C5), 126.2 (d, *J* = 14.0 Hz, benzene C1), 126.1 (benzene C5), 120.8 (d, *J* = 24.2 Hz, benzene C3), 117.3 (tetrahydroquinoline C6), 114.5 (tetrahydroquinoline C8), 46.6 (tetrahydroquinoline C2, ethyl CH_2), 26.9 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 8.1 (CH_3) ppm. EI-MS *m/z* 476 (M^+^{79}Br, 9), 385 (8), 237 (11), 189 (6), 173 (12), 145 (100%). HRMS (ESI) calcd for $C_{17}H_{19}{}^{79}BrFN_2O_4S_2$ (MH+), 476.9959; found, 476.9958. HPLC purity: 97.4%, R_t=6.907 (20% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)thiophene-2-sulfonamide ***175***



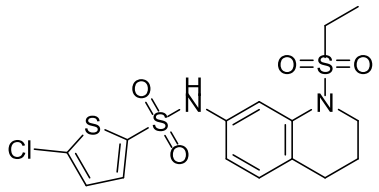
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), thiophene-2-sulfonyl chloride (219 mg), **175** was isolated as a red solid (247 mg, 64%), mp: 80-81 ºC; IR (cm^{-1}): υ 3276 (w, NH), 2942 (m, C-H aliphatic), 1332 (m, SO_2), 1123 (w, C-N). ^{1}H NMR, δ: 7.58 (d, *J* = 3.5 Hz, 1H, thiophene_5 H), 7.54 (d, *J* = 4.9 Hz, 1H, thiophene_3 H), 7.38 (s, 1H, tetrahydroquinoline_8 H), 7.04-7.01 (m, 2H, thiophene_4 H, tetrahydroquinoline_5 H), 6.91 (d, *J* = 7.5 Hz, 1H, tetrahydroquinoline_6 H), 6.65 (s, br, NH), 3.76 (t, *J* = 6.0 Hz, 2H, tetrahydroquinoline_2 H), 3.06 (q, *J* = 7.5 Hz, 2H, ethyl CH_2), 2.78 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.94 (p, *J* = 6.0 Hz, 2H, tetrahydroquinoline_3 H), 1.30 (t, *J* = 7.5 Hz, 3H, CH_3). ^{13}C NMR, δ: 139.4 (thiophene C2), 137.5 (tetrahydroquinoline C8a), 134.9 (tetrahydroquinoline C7), 133.1 (tetrahydroquinoline C5), 132.4 (thiophene C4), 130.5 (thiophene C3), 127.4 (thiophene C5), 125.6 (tetrahydroquinoline C4a), 117.2 (tetrahydroquinoline C6), 114.4 (tetrahydroquinoline C8), 46.5 (tetrahydroquinoline C2), 46.4 (ethyl CH_2), 26.7 (tetrahydroquinoline C4), 22.4 (tetrahydroquinoline C3), 7.9 (CH_3) ppm. EI-MS *m/z* 386 (M+, 20), 322 (20), 293 (11), 229 (29), 145 (96), 118 (100%), 99 (33). HRMS (ESI) calcd for $C_{15}H_{19}N_2O_4S_3$ (MH+), 387.0507; found, 387.0504. HPLC purity: 97.5%, R_t=5.716 (20% solvent D in solvent A).

5-Bromo-N-(1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)thiophene-2-sulfonamide ***176***



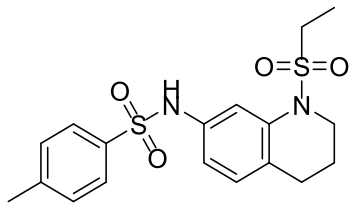
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 5-bromothiophene-2-sulfonyl chloride (313.8 mg), **176** was isolated as red solid (315 mg, 68%), mp: 82-83 ºC; IR (cm^{-1}): υ 3192 (w, NH), 2958 (m, C-H aliphatic), 1326 (m, SO_2), 1152 (w, C-N), 635 (w, C-Br). ^{1}H NMR, δ: 7.41 (d, *J* = 1.5 Hz, 1H, tetrahydroquinoline_8 H), 7.33 (d, *J* = 4.1 Hz, 1H, thiophene_4 H), 7.27 (s, br, NH), 7.04 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.98 (d, *J* = 3.9 Hz, 1H, thiophene_3 H), 6.94 (dd, *J* = 8.4, 1.6 Hz, 1H, tetrahydroquinoline_6 H), 3.77 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.10 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.79 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 1.98 (p, *J* = 5.9 Hz, 2H, tetrahydroquinoline_3 H), 1.31 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 140.2 (thiophene C2), 137.7 (tetrahydroquinoline C8a), 134.5 (tetrahydroquinoline C7), 133.3 (thiophene C4), 130.7 (tetrahydroquinoline C5), 130.5 (thiophene C3), 125.9 (tetrahydroquinoline C4a), 120.3 (tetrahydroquinoline C6), 117.3 (thiophene C5), 114.4 (tetrahydroquinoline C8), 46.5 (tetrahydroquinoline C2, ethyl CH_2), 26.8 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 8.0 (CH_3) ppm. EI-MS *m/z* 466 (M+ ^{79}Br, ^{81}Br, 10), 400 (12), 321 (12), 307 (11), 227 (35), 145 (97), 118 (100%), 91 (32). HRMS (ESI) calcd for $C_{15}H_{18}{}^{79}BrN_2O_4S_3$ (MH+), 464.9612; found, 464.9617. HPLC purity: 98.0%, R_t=3.968 (20% solvent D in solvent A).

5-Chloro-N-(1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl) thiophene-2-sulfonamide ***177***



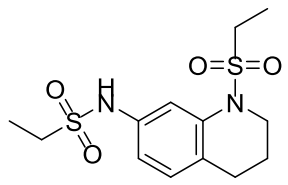
Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 5-chlorothiophene-2-sulfonyl chloride (260.4 mg), **177** was isolated as a yellow oil (277 mg, 66%), IR (cm^{-1}): υ 3195 (w, NH), 2941 (w, C-H aliphatic), 1328 (s, SO_2), 1155 (w, C-N). ^{1}H NMR, δ: 7.48 (s, br, NH), 7.43 (s, 1H, tetrahydroquinoline_8 H), 7.38 (d, *J* = 3.8 Hz, 1H, thiophene_4 H), 7.04 (d, *J* = 8.3 Hz, 1H, tetrahydroquinoline_5 H), 6.94 (d, *J* = 7.9 Hz, 1H, tetrahydroquinoline_6 H), 6.83 (d, *J* = 3.8 Hz, 1H, thiophene_3 H), 3.77 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.12 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.79 (t, *J* = 6.3 Hz, 2H, tetrahydroquinoline_4 H), 1.97 (p, *J* = 5.8 Hz, 2H, tetrahydroquinoline_3 H), 1.30 (t, *J* = 7.3 Hz, 3H, CH_3). ^{13}C NMR, δ: 137.7 (thiophene C2), 137.5 (tetrahydroquinoline C8a), 137.4 (tetrahydroquinoline C7), 134.6 (thiophene C3), 132.7 (thiophene C4), 130.7 (tetrahydroquinoline C5), 126.9 (tetrahydroquinoline C4a), 125.9 (thiophene C5), 117.3 (tetrahydroquinoline C6), 114.4 (tetrahydroquinoline C8), 46.6 (tetrahydroquinoline C2, ethyl CH_2), 26.8 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 8.0 (CH_3) ppm. EI-MS *m/z* 420 (M+ ^{35}Cl, 10), 356 (12), 263 (30), 227 (27), 181 (24), 145 (97), 118 (100%). HRMS (ESI) calcd for $C_{15}H_{18}{}^{35}ClN_2O_4S_3$ (MH+), 421.0117; found, 421.0107. HPLC purity: 99.6%, R_t=7.798 (20% solvent D in solvent A).

N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-4-methylbenzenesulfonamide ***178***



Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 4-methylbenzenesulfonyl chloride (228.7 mg), **178** was isolated as a colourless solid (287 mg, 73%), mp: 74-75 ºC; IR (cm^{-1}): υ 3274 (m, NH), 2939 (w, C-H aliphatic), 1320 (s, SO_2), 1124 (w, C-N). ^{1}H NMR, δ: 7.71 (d, *J* = 7.5 Hz, 2H, benzene_2,6 H), 7.33 (s, 1H, tetrahydroquinoline_8 H), 7.23 (d, *J* = 7.9 Hz, 2H, benzene_3,5 H), 6.98 (d, *J* = 5.5 Hz, 1H, tetrahydroquinoline_5 H), 6.87 (d, *J* = 5.5 Hz, 1H, tetrahydroquinoline_6), 6.85 (s, br, NH), 3.74 (t, *J* = 5.5 Hz, 2H, tetrahydroquinoline_2 H), 3.03 (q, *J* = 7.5 Hz, 2H, ethyl CH_2), 2.75 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 2.37 (s, 3H, *p*-CH_3), 1.94 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.27 (t, *J* = 7.1 Hz, 3H, CH_3). ^{13}C NMR, δ: 144.1 (tetrahydroquinoline C8a), 137.9 (benzene C4), 136.6 (benene C1), 135.5 (tetrahydroquinoline C7), 130.8 (tetrahydroquinoline C5), 130.0 (benzene C3, C5), 127.8 (benzene C2, C6), 125.4 (tetrahydroquinoline C4a), 117.3 (tetrahydroquinoline C6), 114.3 (tetrahydroquinoline C8), 46.8 (tetrahydroquinoline C2), 46.7 (ethyl CH_2), 27.1 (tetrahydroquinoline C4), 22.8 (tetrahydroquinoline C3), 21.9 (benzene_4 CH_3), 8.3 (CH_3) ppm. EI-MS *m/z* 394 (M+, 15), 301 (9), 237 (5), 145 (90), 118 (55), 91 (100%). HRMS (ESI) calcd for $C_{18}H_{23}N_2O_4S_2$ (MH+), 395.1099; found, 395.1096. HPLC purity: 98.9%, R_t=10.216 (20% solvent D in solvent A).

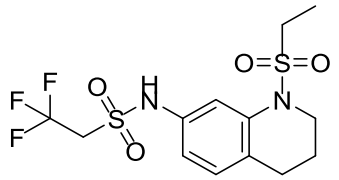
N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)ethanesulfonamide ***179***



Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), ethanesulfonyl chloride (154.2 mg), **179** was isolated

as a colourless oil (199 mg, 60%), IR (cm^{-1}): υ 3260 (m, NH), 2943 (w, C-H aliphatic), 1322 (s, SO_2), 1139 (w, C-N). ^1H NMR, δ: 7.50 (d, *J* = 1.7 Hz, 1H, tetrahydroquinoline_8 H), 7.32 (s, br, NH), 7.06 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.98 (dd, *J* = 8.2, 1.9 Hz, 1H, tetrahydroquinoline_6 H), 3.78 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.21 (q, *J* = 7.4 Hz, 2H, ethylsulfonyl CH_2), 3.15 (q, *J* = 7.4 Hz, 2H, ethanesulfonamide CH_2), 2.80 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 2.00 (p, *J* = 6.4 Hz, 2H, tetrahydroquinoline_3 H), 1.36 (t, *J* = 7.3 Hz, 3H, ethylsulfonyl CH_3), 1.35 (t, *J* = 7.4 Hz, 3H, ethanesulfonamide CH_3). ^{13}C NMR, δ: 137.8 (tetrahydroquinoline C8a), 135.6 (tetrahydroquinoline C7), 130.6 (tetrahydroquinoline C5), 124.9 (tetrahydroquinoline C4a), 116.0 (tetrahydroquinoline C6), 113.5 (tetrahydroquinoline C8), 46.9 (tetrahydroquinoline C2), 46.5 (ethylsulfonyl CH_2), 45.9 (ethanesulfonamide CH_2), 26.6 (tetrahydroquinoline C4), 22.6 (tetrahydroquinoline C3), 8.1 (ethylsulfonyl CH_3), 8.0 (ethanesulfonamide CH_3) ppm. ESI-MS *m/z* 331.1(M-H). HRMS (ESI) calcd for $C_{13}H_{20}N_2O_4S_2$ (M-H), 331.0786; found, 331.0795. HPLC purity: 99.6%, R_t=10.071 (25% solvent D in solvent A).

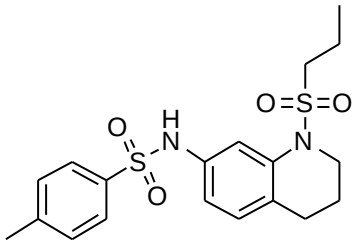
N-(1-(Ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-2,2,2-trifluoroethanesulfonamide ***180***



Following the general procedure 10 using 1-(ethylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **163** (240 mg, 1.0 mmol), 2,2,2-trifluoroethane-1-sulfonyl chloride (219 mg), **180** was isolated as a pale yellow solid (247 mg, 64%), mp: 78-79 ºC; IR (cm^{-1}): υ 3236 (m, NH), 2951 (m, C-H aliphatic), 1244 (w, C-F), 1322 (s, SO_2), 1126 (w, C-N). ^1H NMR, δ: 7.16 (s, 1H, tetrahydroquinoline_8 H), 7.24 (s, br, NH), 7.11 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.99 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 3.85 (q, *J* = 8.8 Hz, 2H, CF_3CH_2), 3.79 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.19 (q, *J* = 7.4 Hz, 2H, ethyl CH_2), 2.83 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 2.02 (p, *J* = 5.5 Hz, 2H, tetrahydroquinoline_3 H), 1.36 (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR, δ: 138.3 (tetrahydroquinoline C8a), 134.1 (tetrahydroquinoline C7), 131.3

(tetrahydroquinoline C5), 126.8 (tetrahydroquinoline C4a), 121.8 (q, *J* = 276.5 Hz, CF_3), 117.3 (tetrahydroquinoline C6), 114.9 (tetrahydroquinoline C8), 52.6 (q, *J* = 31.7 Hz, $CF_3\underline{C}H_2$), 47.2 (tetrahydroquinoline C2), 46.8 (ethyl CH_2), 27.1 (tetrahydroquinoline C4), 22.8 (tetrahydroquinoline C3), 8.3 (CH_3) ppm. ESI-MS *m/z* 385.1 (M-H). HRMS (ESI) calcd for $C_{13}H_{18}F_3N_2O_4S_2$ (MH+), 387.0660; found, 387.0664. HPLC purity: 99.9%, R_t=8.931 (20% solvent D in solvent A).

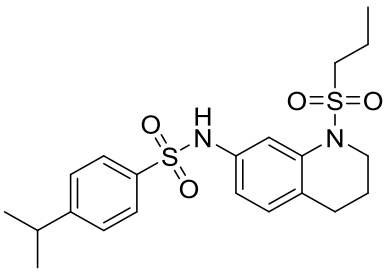
4-Methyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***181***



Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 4-methylbenzenesulfonyl chloride (228.7 mg), **181** was isolated as a pale yellow solid (310 mg, 76%), mp: 112-113 ºC; IR (cm^{-1}): υ 3276 (w, NH), 2976 (m, C-H aliphatic), 1320 (s, SO_2), 1125 (w, C-N). ^{1}H NMR, δ: 7.72 (d, *J* = 7.8 Hz, 2H, benzene_2,6 H), 7.54 (s, br, NH), 7.42 (s, 1H, tetrahydroquinoline_8 H), 7.20 (d, *J* = 7.9 Hz, 2H, benzene_3,5 H), 6.95 (d, *J* = 8.0 Hz, 1H, tetrahydroquinoline_5 H), 6.86 (d, *J* = 8.0 Hz, 1H, tetrahydroquinoline_6 H), 3.72 (t, *J* = 5.6 Hz, 2H, tetrahydroquinoline_2 H), 2.99 (t, *J* = 7.4 Hz, 2H, propyl_1 H), 2.72 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 2.34 (s, 3H, *p*-CH_3), 1.92 (p, *J* = 5.7 Hz, 2H, tetrahydroquinoline_3 H), 3.74 (sex, *J* = 7.4 Hz, 2H, propyl_2 H), 0.97 (t, *J* = 7.3 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 143.9 (tetrahydroquinoline C8a), 137.8 (benzene C4), 136.5 (benzene C1), 135.6 (tetrahydroquinoline C7), 130.6 (tetrahydroquinoline C5), 129.8 (benzene C3, C5), 127.6 (benzene C2, C6), 125.1 (tetrahydroquinoline C4a), 116.9 (tetrahydroquinoline C6), 114.0 (tetrahydroquinoline C8), 53.9 (tetrahydroquinoline C2), 46.6 (propyl C1), 26.9 (tetrahydroquinoline C4), 22.7 (tetrahydroquinoline C3), 21.7 (benzene_4 CH_3), 17.2 (propyl C2), 13.2 (propyl C3) ppm. EI-MS *m/z* 408 (M+, 48), 301 (46), 237 (10), 145 (100%), 118 (45), 91 (81).

HRMS (ESI) calcd for $C_{19}H_{25}N_2O_4S_2$ (MH+), 409.1256; found, 409.1258. HPLC purity: 99.9%, R_t=6.809 (20% solvent D in solvent A).

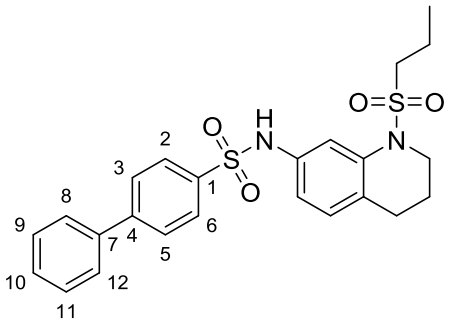
4-Isopropyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***182***



Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 4-isopropylbenzenesulfonyl chloride (555.7 mg), **182** was isolated as pale yellow solid (344 mg, 79%), mp: 108-109 ºC; IR (cm^{-1}): υ 3273 (m, NH), 2967 (m, C-H aliphatic), 1324 (s, SO_2), 1126 (w, C-N). ^{1}H NMR, δ: 7.77 (d, *J* = 8.2 Hz, 2H, benzene_2,6 H), 7.42 (s, 1H, tetrahydroquinoline_8 H), 7.28 (d, *J* = 8.0 Hz, 2H, benzene_3,5 H), 7.21 (s, br, NH), 6.96 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.85 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 3.74 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.02 (t, *J* = 7.6 Hz, 2H, propyl_1 H), 2.95-2.89 (m, 1H, C<u>H</u>$(CH_3)_2$), 2.74 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.94 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.77 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 1.22 (d, *J* = 6.8 Hz, 6H, CH(C<u>H</u>$_3)_2$), 0.98 (t, *J* = 7.5 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 154.5 (benzene C4), 137.7 (tetrahydroquinoline C8a), 136.8 (benzene C1), 135.5 (tetrahydroquinoline C7), 130.5 (tetrahydroquinoline C5), 127.6 (benzene C3, C5), 127.2 (benzene C2, C6), 124.9 (tetrahydroquinoline C4a), 116.6 (tetrahydroquinoline C6), 113.8 (tetrahydroquinoline C8), 53.9 (tetrahydroquinoline C2), 46.5 (propyl C1), 34.2 (<u>C</u>H$(CH_3)_2$), 26.8 (tetrahydroquinoline C4), 23.7 (CH(<u>C</u>$H_3)_2$), 22.6 (tetrahydroquinoline C3), 17.1 (propyl C2), 13.1 (propyl C3) ppm. EI-MS *m/z* 436 (M+, 33), 329 (30), 265 (8), 145 (100%), 119 (60), 91 (47). HRMS (ESI) calcd for $C_{21}H_{29}N_2O_4S_2$ (MH+), 437.1569; found, 437.1550. HPLC purity: 99.4%, R_t=5.180 (15% solvent D in solvent A).

N-(1-(Propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)-[1,1'-biphenyl]-4-sulfonamide[*]

183

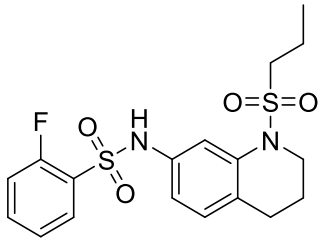


Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), [1,1'-biphenyl]-4-sulfonyl chloride (303.2 mg), **183** was isolated as a reddish white solid (362 mg, 77%), mp: 118-119 ºC; IR (cm^{-1}): υ 3265 (m, NH), 2971 (w, C-H aliphatic), 1325 (s, SO_2), 1125 (w, C-N). ^{1}H NMR, δ: 7.90 (d, *J* = 8.3 Hz, 2H, H2, H6), 7.57 (d, *J* = 8.3 Hz, 2H, H3, H5), 7.50 (d, *J* = 7.5 Hz, 2H, H8, H12), 7.46 (d, *J* = 1.3 Hz, 1H, tetrahydroquinoline_8 H), 7.42 (s, br, NH), 7.40 (d†, *J* = 7.7 Hz, 2H, H9, H11), 7.36 (t, *J* = 7.0 Hz, 1H, H10), 6.97 (d, *J* = 8.3 Hz, 1H, tetrahydroquinoline_5 H), 6.90 (dd, *J* = 8.2, 1.4 Hz, 1H, tetrahydroquinoline_6 H), 3.71 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline_2 H), 3.00 (t, *J* = 7.6 Hz, 2H, propyl _1 H), 2.70 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.91 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.72 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 0.90 (t, *J* = 7.5 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 145.9 (C7), 139.4 (C4), 138.2 (tetrahydroquinoline C8a), 138.0 (C1), 135.6 (tetrahydroquinoline C7), 130.7 (C3, C5), 129.2 (tetrahydroquinoline C5), 128.7 (C2, C6), 128.2 (C9, C11), 127.8 (tetrahydroquinoline C4a), 127.5 (C10), 125.5 (C8, C12), 117.1 (tetrahydroquinoline C6), 114.4 (tetrahydroquinoline C8), 54.1 (tetrahydroquinoline C2), 46.7 (propyl C1), 27.0 (tetrahydroquinoline C4), 22.8 (tetrahydroquinoline C3), 17.3 (propyl C2), 13.1 (propyl C3) ppm. EI-MS *m/z* 470 (M+, 45), 363 (33), 299 (29), 253 (10), 207 (11), 145 (100%), 118 (62), 91 (31). HRMS (ESI) calcd for $C_{24}H_{27}N_2O_4S_2$ (MH+), 471.1412; found, 471.1407. HPLC purity: 98.8%, R_t=6.716 (20% solvent D in solvent A).

[*] Numbering of the 1,1'-biphenyl carbons doesn't correspond to the compound name.
† Apparent doublet.

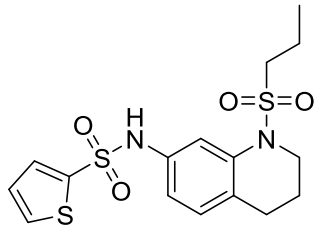
2-Fluoro-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide

184



Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 2-fluorobenzenesulfonyl chloride (233.5 mg), **184** was isolated as red solid (284 mg, 69%), mp: 102-103 ºC; IR (cm^{-1}): υ 3245 (w, NH), 2958 (m, C-H aliphatic), 1472 (w, C-F), 1331 (s, SO_2), 1126 (w, C-N). ^{1}H NMR, δ: 7.55 (t, *J* = 7.6 Hz, 1H, benzene_4 H), 7.53 (q, *J* = 7.4 Hz, 1H, benzene_3 H), 7.42 (s, 1H, tetrahydroquinoline_8 H), 7.23-7.17 (m, 2H, benzene_5,6 H), 6.97 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.94 (s, br, NH), 6.86 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_6 H), 3.72 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 2.95 (t, *J* = 7.7 Hz, 2H, propyl_1 H), 2.73 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.92 (p, *J* = 6.2 Hz, 2H, tetrahydroquinoline_3 H), 1.75 (sex, *J* = 7.7 Hz, 2H, propyl_2 H), 0.99 (t, *J* = 7.6 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 159.0 (d, *J* = 254.96 Hz, benzene C2), 137.9 (tetrahydroquinoline C8a), 135.6 (d, *J* = 8.38 Hz, benzene C4), 134.6 (tetrahydroquinoline C7), 131.3 (tetrahydroquinoline C5), 130.7 (tetrahydroquinoline C4a), 127.1 (d, *J* = 13.07 Hz, benzene C1), 125.9 (tetrahydroquinoline C6), 124.7 (d, *J* = 3.77 Hz, benzene C6), 117.3 (benzene C5), 117.1 (d, *J* = 20.47 Hz, benzene C3), 114.5 (tetrahydroquinoline C8), 53.9 (tetrahydroquinoline C2), 46.6 (propyl C1), 27.0 (tetrahydroquinoline C4), 22.6 (tetrahydroquinoline C3), 17.2 (propyl C2), 13.2 (propyl C3) ppm. EI-MS *m/z* 412 (M+, 42), 305 (44), 145 (100%), 118 (60). HRMS (ESI) calcd for $C_{18}H_{22}FN_2O_4S_2$ (MH+), 413.1005; found, 413.1000. HPLC purity: 99.7%, R_t=8.117 (20% solvent D in solvent A).

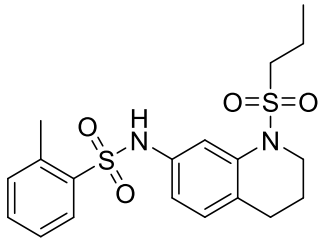
N-(1-(Propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)thiophene-2-sulfonamide ***185***



Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), thiophene-2-sulfonyl chloride (219 mg), **185** was isolated as an orange solid (268 mg, 67%), mp: 96-97 ºC; IR (cm^{-1}): υ 3288 (w, NH), 2976 (w, C-H aliphatic), 1335 (s, SO_2), 1125 (w, C-N). ^{1}H NMR, δ: 7.58 (d, *J* = 3.7 Hz, 1H, thiophene_5 H), 7.52 (d, *J* = 4.9 Hz, 1H, thiophene_3 H), 7.43 (s, 1H, tetrahydroquinoline_8 H), 7.18 (s, br, NH), 7.03 (d, *J* = 7.9 Hz, 1H, tetrahydroquinoline_5 H), 7.00 (t, *J* = 4.2 Hz, 1H, thiophene_4), 6.92 (d, *J* = 8.1 Hz, 1H, tetrahydroquinoline_6), 3.75 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 3.02 (t, *J* = 7.8 Hz, 2H, propyl_1 H), 2.77 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 1.96 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.78 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 1.00 (t, *J* = 7.5 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 139.6 (thiophene C2), 137.6 (tetrahydroquinoline C8a), 134.8 (tetrahydroquinoline C7), 133.1 (thiophene C4), 132.4 (tetrahydroquinoline C5), 130.5 (thiophene C3), 127.4 (thiophene C5), 125.6 (tetrahydroquinoline C4a), 117.3 (tetrahydroquinoline C6), 114.5 (tetrahydroquinoline C8), 53.7 (tetrahydroquinoline C2), 46.4 (propyl C1), 26.8 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 17.0 (propyl C2), 12.9 (propyl C3) ppm. EI-MS *m/z* 400 (M+, 24), 335 (25), 293 (17), 229 (61), 145 (100%), 118 (76). HRMS (ESI) calcd for $C_{16}H_{21}N_2O_4S_3$ (MH+), 401.0663; found, 401.0646. HPLC purity: 99.9%, R_t=8.586 (20% solvent D in solvent A).

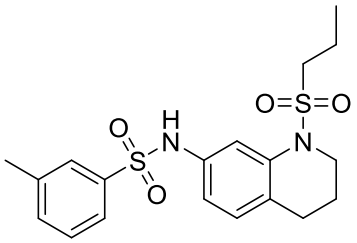
2-Methyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide

186



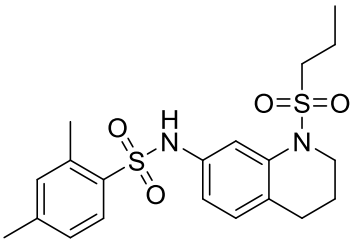
Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 2-methylbenzenesulfonyl chloride (228.7 mg), **186** was isolated as an orange oil (302 mg, 74%), IR (cm^{-1}): υ 3244 (m, NH), 2936 (m, C-H aliphatic), 1326 (s, SO_2), 1130 (w, C-N). ^{1}H NMR, δ: 7.98 (d, *J* = 7.7 Hz, 1H, benzene_6 H), 7.41 (t, *J* = 7.7 Hz, 1H, benzene_5 H), 7.35 (d, *J* = 2.0 Hz, 1H, tetrahydroquinoline_8 H), 7.27-7.24 (m, 3H, NH, benzene_3,4 H), 6.93 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.81 (dd, *J* = 8.2, 2.0 Hz, 1H, tetrahydroquinoline_6 H), 3.70 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline_2 H), 2.94 (t, *J* = 7.9 Hz, 2H, propyl_1 H), 2.70 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 2.65 (s, 3H, *o*-CH_3), 1.90 (p, *J* = 6.2 Hz, 2H, tetrahydroquinoline_3 H), 1.72 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 0.96 (t, *J* = 7.4 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 137.8 (benzene C1), 137.7 (tetrahydroquinoline C8a), 137.5 (benzene C2), 135.4 (benzene C4), 133.2 (benzene C5), 132.8 (tetrahydroquinoline C7), 130.7 (benzene C3), 130.4 (tetrahydroquinoline C5), 126.4 (benzene C6), 124.9 (tetrahydroquinoline C4a), 116.2 (tetrahydroquinoline C6), 113.5 (tetrahydroquinoline C8), 53.8 (tetrahydroquinoline C2), 46.6 (propyl C1), 26.9 (tetrahydroquinoline 4), 22.7 (tetrahydroquinoline C3), 20.6 (benzene_2 CH_3), 17.2 (propyl C2), 13.2 (propyl C3) ppm. EI-MS *m/z* 408 (M+, 40), 301 (36), 237 (10), 145 (96), 118 (55), 91 (100%). HRMS (ESI) calcd for $C_{19}H_{25}N_2O_4S_2$ (MH+), 409.1256; found, 409.1252. HPLC purity: 99.7%, R_t=6.298 (20% solvent D in solvent A).

3-Methyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***187***



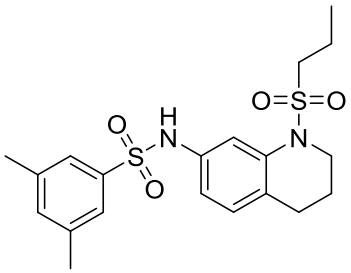
Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 3-methylbenzenesulfonyl chloride (228.7 mg), **187** was isolated as a yellow oil (306 mg, 75%), IR (cm^{-1}): υ 3256 (m, NH), 2934 (m, C-H aliphatic), 1324 (s, SO_2), 1123 (m, C-N). ^{1}H NMR, δ: 7.69 (s, 1H, benzene_2 H), 7.61 (d, *J* = 6.4 Hz, 1H, benzene_6 H), 7.38 (s, 1H, tetrahydroquinoline_8 H) 7.31-7.28 (m, 2H, benzene_4,5 H), 7.22 (s, br, NH), 6.96 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.86 (dd, *J* = 8.2, 1.8 Hz, 1H, tetrahydroquinoline_6 H), 3.72 (t, *J* = 5.9 Hz, 2H, tetrahydroquinoline_2 H), 2.98 (t, *J* = 7.6 Hz, 2H, propyl_1 H), 2.73 (t, *J* = 6.6 Hz, 2H, tetrahydroquinoline_4 H), 2.34 (s, 3H, *m*-CH_3), 1.92 (p, *J* = 6.2 Hz, 2H, tetrahydroquinoline_3 H), 1.74 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 0.97 (t, *J* = 7.6 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 139.3 (tetrahydroquinoline C8a), 139.0 (benzene C3), 137.5 (benzene C1), 135.3 (tetrahydroquinoline C7), 133.7 (benzene C4), 130.4 (tetrahydroquinoline C5), 128.8 (benzene C5), 127.7 (tetrahydroquinoline C4a), 124.9 (benzene C6), 124.5 (benzene C2), 116.8 (tetrahydroquinoline C6), 113.9 (tetrahydroquinoline C8), 53.7 (tetrahydroquinoline C2), 46.4 (propyl C1), 26.7 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 21.2 (benzene_3 CH_3), 17.0 (propyl C2), 12.9 (propyl C3) ppm. EI-MS *m/z* 408 (M+, 43), 301 (37), 237 (11), 145 (100%), 118 (60), 91 (95). HRMS (ESI) calcd for $C_{19}H_{25}N_2O_4S_2$ (MH+), 409.1256; found, 409.1236. HPLC purity: 99.5%, R_t=6.382 (20% solvent D in solvent A).

2,4-Dimethyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***188***



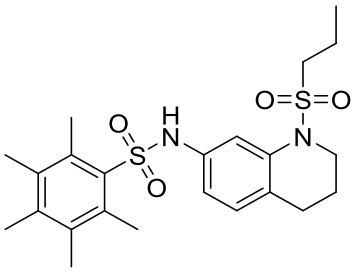
Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 2,4-dimethylbenzenesulfonyl chloride (245.5 mg), **188** was isolated as an orange oil (320 mg, 76%), IR (cm^{-1}): υ 3277 (m, NH), 2973 (m, C-H aliphatic), 1322 (s, SO_2), 1139 (m, C-N). ^{1}H NMR, δ: 7.88 (d, *J* = 8.0 Hz, 1H, benzene_6 H), 7.35 (d, *J* = 2.0 Hz 1H, tetrahydroquinoline_8 H), 7.33 (s, 1H, benzene_3 H), 7.06 (s, br, NH), 7.05 (d, *J* = 8.0 Hz, 1H, benzene_5 H), 6.92 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.80 (dd, *J* = 8.5, 2.0 Hz, 1H, tetrahydroquinoline_6 H), 3.72 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline_2 H), 2.95 (t, *J* = 7.7 Hz, 2H, propyl_1 H), 2.71 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 2.61 (s, 3H, benzene_2 CH_3), 2.31 (s, 3H, benzene_4 CH_3), 1.91 (p, *J* = 6.0 Hz, 2H, tetrahydroquinoline_3 H), 1.75 (sex, *J* = 7.7 Hz, 2H, propyl_2 H), 0.97 (t, *J* = 7.5 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 143.9 (benzene C1), 137.8 (benzene C2), 137.3 (tetrahydroquinoline C8a), 135.6 (benzene C4), 134.8 (benzene C5), 133.5 (tetrahydroquinoline C7), 130.6 (benzene C6), 130.6 (tetrahydroquinoline C5), 127.0 (tetrahydroquinoline C4a), 124.7 (benzene C3), 116.0 (tetrahydroquinoline C6), 113.3 (tetrahydroquinoline C8), 53.8 (tetrahydroquinoline C2), 46.6 (propyl C1), 26.9 (tetrahydroquinoline C4), 22.7 (tetrahydroquinoline C3), 21.5 (benzene_2 CH_3), 20.5 (benzene_4 CH_3), 17.2 (propyl C2), 13.2 (propyl C3) ppm. EI-MS *m/z* 422 (M+, 32), 315 (26), 251 (29), 145 (100%), 118 (33), 105 (30), 91 (17). HRMS (ESI) calcd for $C_{20}H_{27}N_2O_4S_2$ (MH+), 423.1412; found, 423.1399. HPLC purity: 99.9%, R_t=5.799 (20% solvent D in solvent A).

3,5-Dimethyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***189***



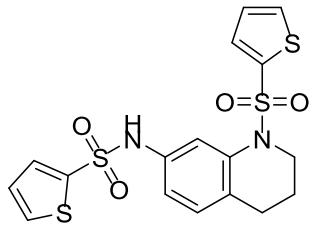
Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 3,5-dimethylbenzenesulfonyl chloride (245.5 mg), **189** was isolated as an orange oil (324 mg, 77%), IR (cm^{-1}): υ 3246 (m, NH), 2968 (m, C-H aliphatic), 1324 (s, SO_2), 1146 (w, C-N). ^{1}H NMR, δ: 7.48 (s, 2H, benzene_2,6 H), 7.42-7.40 (m, 2H, tetrahydroquinoline_8 H, benzene_4 H), 7.11 (s, br, NH), 6.95 (d, *J* = 8.3 Hz, 1H, tetrahydroquinoline_5 H), 6.85 (dd, *J* = 8.2, 1.70 Hz, 1H, tetrahydroquinoline_6 H), 3.72 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline_2 H), 2.99 (t, *J* = 7.7 Hz, 2H, propyl_1 H), 2.73 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 2.29 (s, 6H, benzene_3,5 CH_3), 1.92 (p, *J* = 6.0 Hz, 2H, tetrahydroquinoline_3 H), 1.75 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 0.97 (t, *J* = 7.5 Hz, 3H, propyl_1 H). ^{13}C NMR, δ: 138.9 (tetrahydroquinoline C8a), 138.8 (benzene C3, C5), 137.5 (benzene C1), 135.3 (tetrahydroquinoline C7), 134.5 (benzene C4), 130.3 (tetrahydroquinoline C5), 124.9 (tetrahydroquinoline C4a), 124.7 (benzene C2, C6), 116.5 (tetrahydroquinoline C6), 113.6 (tetrahydroquinoline C8), 53.6 (tetrahydroquinoline C2), 46.3 (propyl C1), 26.6 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 21.1 (benzene_3,5 CH_3), 16.9 (propyl C2), 12.9 (propyl C3) ppm. EI-MS *m/z* 422 (M+, 40), 315 (38), 251 (20), 207 (17), 145 (100%), 118 (27), 105 (29). HRMS (ESI) calcd for $C_{20}H_{27}N_2O_4S_2$ (MH+), 423.1412; found, 423.1396. HPLC purity: 99.9%, R_t=5.503 (20% solvent D in solvent A).

2,3,4,5,6-Pentamethyl-N-(1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl) benzenesulfonamide ***190***



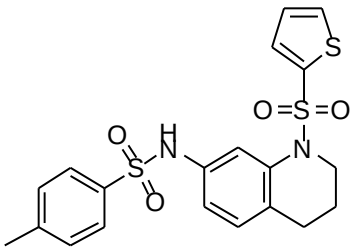
Following the general procedure 10 using 1-(propylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **167** (254 mg, 1.0 mmol), 2,3,4,5,6-pentamethylbenzenesulfonyl chloride (296 mg), **190** was isolated as a yellow solid (371 mg, 80%), mp: 128-129 ºC; IR (cm^{-1}): υ 3266 (m, NH), 2937 (m, C-H aliphatic), 1315 (s, SO_2), 1130 (w, C-N). ^{1}H NMR, δ: 7.21 (s, 1H, tetrahydroquinoline_8 H), 6.92 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.92 (s, br, NH), 6.70 (d, *J* = 8.2, 1H, tetrahydroquinoline_6 H), 3.71 (t, *J* = 5.7 Hz, 2H, tetrahydroquinoline_2 H), 2.84 (t, *J* = 7.8 Hz, 2H, propyl_1 H), 2.71 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 2.58 (s, 6H, benzene_2,6 CH_3), 2.24 (s, 3H, benzene_4 CH_3), 2.21 (s, 6H, benzene_3,5 CH_3), 1.91 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H), 1.67 (sex, *J* = 7.6 Hz, 2H, propyl_2 H), 0.93 (t, *J* = 7.4 Hz, 3H, propyl_3 H). ^{13}C NMR, δ: 139.8 (benzene C2, C6), 137.5 (tetrahydroquinoline C8a), 135.9 (benzene C3, C5), 135.4 (benzene C4), 134.9 (tetrahydroquinoline C7), 134.4 (tetrahydroquinoline C5), 130.4 (tetrahydroquinoline C4a), 124.6 (benzene C1), 116.4 (tetrahydroquinoline C6), 113.6 (tetrahydroquinoline C8), 53.4 (tetrahydroquinoline C2), 46.4 (propyl C1), 26.8 (tetrahydroquinoline C4), 22.5 (tetrahydroquinoline C3), 19.0 (benzene_2,6 CH_3), 17.8 (benzene_3,5 CH_3), 17.1 (benzene_4 CH_3), 16.9 (propyl C2), 12.9 (propyl C3) ppm. EI-MS *m/z* 464 (M+, 29), 293 (66), 254 (16), 147 (100%). HRMS calcd for $C_{23}H_{33}N_2O_4S_2$ (MH+), 465.1882; found, 465.1881. HPLC purity: 99.9%, R_t=5.024 (15% solvent D in solvent A).

N-(1-(Thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl)thiophene-2-sulfonamide ***191***



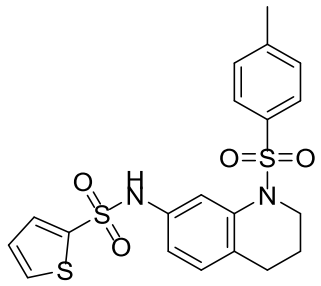
Following the general procedure 10 using 1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **168** (294 mg, 1.0 mmol), thiophene-2-sulfonyl chloride (219 mg), **191** was isolated as a white solid (264 mg, 60%), mp: 108-109 ºC; IR (cm^{-1}): υ 3259 (m, NH), 2965 (w, C-H aliphatic), 1339 (s, SO_2), 1122 (w, C-N). ^{1}H NMR, δ: 7.60 (d, *J* = 3.4 Hz, 1H, thiophen-2-yl_5 H), 7.55 (d, *J* = 4.9 Hz, 1H, thiophene_5 H), 7.50 (d, *J* = 4.9 Hz, 2H, thiophen-2-yl_3 H, thiophene_3 H), 7.35 (d, *J* = 3.3 Hz, 1H, thiophen-2-yl_4 H), 7.26 (s, br, NH), 7.06-6.96 (m, 3H, tetrahydroquinoline_5,6 H, thiophene_4 H), 6.76 (s, 1H, tetrahydroquinoline_8H), 3.79 (t, *J* = 5.8 Hz, 2H, tetrahydroquinoline _2 H), 2.46 (t, *J* = 6.5 Hz, 2H, tetrahydroquinoline_4 H), 1.66 (p, *J* = 6.0 Hz, 2H, tetrahydroquinoline _3 H). ^{13}C NMR, δ: 139.5 (tetrahydroquinoline C8a), 139.2 (thiophene C1), 136.9 (thiophen-2-yl C1), 134.6 (tetrahydroquinoline C7), 133.3 (tetrahydroquinoline C5), 132.5 (thiophene C4), 132.4 (thiophen-2-yl C3), 132.3 (thiophene C3), 130.1 (thiophen-2-yl C4), 128.0 (thiophen-2-yl C5), 127.5 (thiophene C5), 127.4 (tetrahydroquinoline C4a), 118.9 (tetrahydroquinoline C6), 117.7 (tetrahydroquinoline C8), 46.8 (tetrahydroquinoline C2), 26.4 (tetrahydroquinoline C4), 21.3 (tetrahydroquinoline C5) ppm. EI-MS *m/z* 440 (M+, 29), 376 (26), 312 (38), 229 (80), 145 (100%), 118 (78), 91 (55). HRMS (ESI) calcd for $C_{17}H_{17}N_2O_4S_4$ (MH+), 441.0071; found, 441.0068. HPLC purity: 99.9%, R_t=9.126 (20% solvent D in solvent A).

4-Methyl-N-(1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-yl) benzenesulfonamide ***192***



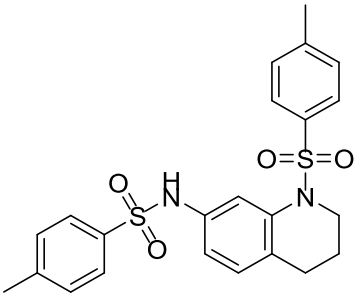
Following the general procedure 10 using 1-(thiophen-2-ylsulfonyl)-1,2,3,4-tetrahydroquinolin-7-amine **168** (294 mg, 1.0 mmol), 4-methylbenzenesulfonyl chloride (228.7 mg), **192** was isolated as a yellow solid (291 mg, 65%), mp: 119-120 ºC; IR (cm^{-1}): υ 3236 (m, NH), 2948 (w, C-H aliphatic), 1318 (s, SO_2), 1129 (w, C-N). ^{1}H NMR, δ: 7.74 (d, *J* = 8.1 Hz, 2H, benzene_2,6 H), 7.50 (s, 1H, tetrahydroquinoline_8 H), 7.48 (d, *J* = 4.8 Hz, 1H, thiophene_5 H), 7.28 (d, *J* = 3.1 Hz, 1H, thiophene_3 H), 7.23 (d, *J* = 8.0 Hz, 2H, benzene_3,5 H), 6.97-6.91 (m, 3H, tetrahydroquinoline_5,6 H, thiophene_4 H), 6.89 (s, br, NH), 3.77 (t, *J* = 6.1 Hz, 2H, tetrahydroquinoline_2 H), 2.43 (t, *J* = 6.4 Hz, 2H, tetrahydroquinoline_4 H), 2.36 (s, 3H, *p*-CH_3), 1.65 (p, *J* = 6.1 Hz, 2H, tetrahydroquinoline_3 H). ^{13}C NMR, δ: 143.8 (tetrahydroquinoline C8a), 139.2 (thiophene C2), 136.8 (benzene C4), 136.2 (benzene C1), 135.0 (tetrahydroquinoline C7), 132.3 (tetrahydroquinoline C5), 132.1 (thiophene C3), 129.9 (thiophene C4), 129.5 (thiophene C5), 127.3 (benzene C3, C5), 127.3 (tetrahydroquinoline C4a), 127.2 (benzene C2, C6), 118.3 (tetrahydroquinoline C6), 117.2 (tetrahydroquinoline C8), 46.8 (tetrahydroquinoline C2), 26.3 (tetrahydroquinoline C4), 21.6 (tetrahydroquinoline C3), 21.3 (benzene_4 CH_3) ppm. EI-MS *m/z* 448 (M+, 24), 384 (17), 319 (26), 229 (99), 146 (100%), 118 (53), 91 (80). HRMS (ESI) calcd for $C_{20}H_{21}N_2O_4S_3$ (MH+), 449.0663; found, 449.0645. HPLC purity: 99.8%, R_t=7.260 (20% solvent D in solvent A).

N-(1-Tosyl-1,2,3,4-tetrahydroquinolin-7-yl)thiophene-2-sulfonamide ***193***



Following the general procedure 10 using 1-tosyl-1,2,3,4-tetrahydroquinolin-7-amine **169** (302 mg, 1.0 mmol), thiophene-2-sulfonyl chloride (219 mg), **193** was isolated as a pale yellow solid (287 mg, 64%), mp: 68-69 ºC; IR (cm^{-1}): υ 3255 (m, NH), 2951 (w, C-H aliphatic), 1329 (s, SO_2), 1132 (w, C-N). ^{1}H NMR, δ: 7.60 (d, *J* = 3.8 Hz, 1H, thiophene_5 H), 7.55 (d, *J* = 4.9 Hz, 1H, thiophene_3 H), 7.51 (d, *J* = 1.7 Hz, 1H, tetrahydroquinoline_8 H), 7.41 (d, *J* = 8.0 Hz, 2H, tosyl_2,6 H), 7.18 (d, *J* = 8.1 Hz, 2H, tosyl_3,5 H), 7.02 (t, *J* = 4.0 Hz, 1H, thiophene_4 H), 7.00 (dd, *J* = 7.9, 1.7 Hz, 1H, tetrahydroquinoline_6 H), 6.92 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 6.83 (s, br, NH), 3.75 (t, *J* = 5.9 Hz, 2H, tetrahydroquinoline_2 H), 2.42 (t, *J* = 6.6 Hz, 2H, tetrahydroquinoline_4 H), 2.38 (s, 3H, *p*-CH_3), 1.59 (p, *J* = 6.2 Hz, 2H, tetrahydroquinoline_3 H). ^{13}C NMR, δ: 143.8 (thiophene C2), 139.5 (tetrahydroquinoline C8a), 137.5 (tosyl C4), 136.5 (tetrahydroquinoline C7), 134.5 (tosyl C1), 133.3 (tetrahydroquinoline C5), 132.4 (thiophene C4), 129.9 (thiophene C3), 129.7 (tosyl C3, C5), 127.6 (thiophene C5), 127.4 (tosyl C2, C6), 127.1 (tetrahydroquinoline C4a), 118.5 (tetrahydroquinoline C6), 117.5 (tetrahydroquinoline C8), 46.5 (tetrahydroquinoline C2), 26.3 (tetrahydroquinoline C4), 21.6 (tetrahydroquinoline C3), 21.3 (CH_3) ppm. EI-MS *m/z* 448 (M+, 16), 320 (19), 229 (18), 207 (22), 145 (100%) 118 (60), 91 (85). HRMS (ESI) calcd for $C_{20}H_{21}N_2O_4S_3$ (MH+), 449.0663; found, 449.0655. HPLC purity: 98.9%, R_t=19.993 (15% solvent D in solvent A).

4-Methyl-N-(1-tosyl-1,2,3,4-tetrahydroquinolin-7-yl)benzenesulfonamide ***194***



Following the general procedure 10 using 1-tosyl-1,2,3,4-tetrahydroquinolin-7-amine **169** (302 mg, 1 mmol), 4-methylbenzenesulfonyl chloride (228.7 mg), **194** was isolated as a white solid (296 mg, 65%), mp: 148-149 ºC; IR (cm^{-1}): υ 3232 (m, NH), 2921 (w, C-H aliphatic), 1326 (s, SO_2), 1157 (w, C-N). ^{1}H NMR, δ: 7.75 (d, *J* = 8.1 Hz, 2H, benzene_2,6 H), 7.51 (d, *J* = 2.1 Hz, 1H, tetrahydroquinoline_8 H), 7.37 (d, *J* = 8.1 Hz, 2H, benzene_3,5 H), 7.23 (d, *J* = 8.1 Hz, 2H, tosyl_2,6 H), 7.14 (d, *J* = 8.1 Hz, 2H, tosyl_3,5 H), 7.01 (s, br, NH), 6.93 (dd, *J* = 8.0, 1.9 Hz, 1H, tetrahydroquinoline_6 H), 6.86 (d, *J* = 8.2 Hz, 1H, tetrahydroquinoline_5 H), 3.73 (t, *J* = 5.9 Hz, 2H, tetrahydroquinoline_2 H), 2.39-2.36 (m, 5H, tetrahydroquinoline_4 H, benzene_4 CH_3), 2.35 (s, 3H, tosyl_CH_3), 1.57 (p, *J* = 5.9 Hz, 2H, tetrahydroquinoline_3 H). ^{13}C NMR, δ: 143.7, 143.6, 137.4, 136.5, 136.2, 134.9, 129.9 (tetrahydroquinoline C5), 129.7, 129.6, 127.5, 127.1, 126.9 (tetrahydroquinoline C6), 117.9 (tetrahydroquinoline C8), 116.9, 46.5 (tetrahydroquinoline C2), 26.2 (tetrahydroquinoline C4), 21.6 (tetrahydroquinoline C3), 21.5 (benzene_4 CH_3), 21.3 (tosyl_CH_3) ppm. EI-MS *m/z* 456 (M+, 30), 327 (27), 301 (33), 237 (24), 207 (19), 145 (100%), 118 (37), 91 (73). HRMS (ESI) calcd for $C_{23}H_{25}N_2O_4S_2$ (MH+), 457.1256; found, 457.1242. HPLC purity: 99.4%, R_t=13.965 (15% solvent D in solvent A).

9.2. *In Silico* experiments

9.2.1. CHIKV non-structural protein 2 (nsP2)

The CHIKV nsP2 crystal structure (PDB: 3TRK) was used for the virtual screening study. The virtual screening study was performed on both nsP2 sites (the domain C site_1 and the domain N site_2) using the Life chemicals cysteine protease inhibitors library (28,960 compounds). This cysteine protease inhibitors library was designed

using Ligand based approach – first, a set of 585 compounds active in assays related to cysteine proteases was assembled. And then, Life Chemicals collection was searched for compounds similar to the reference dataset using MDL public keys and the Tanimoto similarity cut-off of 85% to generate a library of 28,960 compounds. Omega2 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com),[235, 236] was used to generate multiple conformers for each compound in the cysteine protease inhibitors library, using the default settings, generating 3,349,162 conformers. The binding sites were prepared for docking using Fred receptor setup software (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com). The binding sites were detected using the molecular detection method within Fred receptor setup software, where the method detection uses multiple molecular probes rather than a single carbon probe to determine areas of the protein where docking is likely to occur. These probes are docked using a shape based potential, and regions where multiple probes dock are considered favourable. This type of site based detection is slow, it can take several minutes, but detects few sites that are in general of higher quality than atomic probe detection method.

For the C domain site_1, the binding site detected within the C domain had a size of 8321 $Å^3$, an inner contour of 119 $Å^3$ and an outer contour of 2023 $Å^3$. For the N domain site_2, the binding site had a box size of 9394 $Å^3$, an inner contour of 69 $Å^3$ and an outer contour of 2748 $Å^3$. As a blind docking, no constraints were enabled within these sites, with any of the amino acid residues.

Fast exhaustive virtual screening was performed using FRED v2.2.5 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com). During the docking calculations, chemgauss3 scoring functions was enabled. After the docking calculations, the poses returned were scored and ranked with a Gaussian shape function independently by the five available scoring functions (PLP, Chemgauss3, Chemscore, OEChemscore, and Screenscore) and by a consensus of all. The top ranked poses from the exhaustive docking were then optimized using systematic solid body optimization by chemgauss3. VIDA v4.2.0 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com) was used to visualise the docked poses within the receptor active site, and to inspect the critical interacting residues in each pocket with

the individual docked poses. Top 25 hits were then recorded for each of the two sites (Tables 5.1, 5.4).

The top 25 docked poses ranked in each of the two binding sites (site_1, site_2) were then extracted as PDB files, and were processed with AutoDock Tools 1.5.6rc3 (ADT) graphical interface.[173] The Gasteiger charges were computed and the nonpolar hydrogen atoms were merged, torsion angles were defined, they were then saved as pdbqt files for Autodock calculations. The crystal structure 3TRK was used by ADT to setup the binding sites. For the C domain site_1, the grid box x, y and z coordinates were 11.779, 31.511 and 29.546, respectively. The grid spacing was set to 0.375 Å. The grid box size was set to 42 x 42 x 42 points in x, y, and a z direction. For the N domain site_2, the grid box x, y and z coordinates were 3.366, 28.678 and 21.482 respectively. The grid box size was set to 42 x 42 x 42 points in x, y, and a z direction, with grid spacing of 0.375 Å. AutoGrid 4.2 algorithm was used to evaluate the binding energies between the inhibitors and the enzyme and to generate the energy maps for the docking run. Fifty runs were generated by using Autodock 4.2 Lamarckian genetic algorithm[173] for the searches. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 1.0, 2.0 and 3.0 Å, the docked poses were ranked according to the binding energies and ligand efficiencies, and finally the five lowest energy poses **(**Tables 5.2, 5.3, 5.5, 5.6**)** were selected as the resultant complexes with the enzymes. The complexes were then typed with the CHARMm forcefield with Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012) to relax the obtained poses within the enzyme pockets, and visualized.

9.2.2. CHIKV envelope proteins

9.2.2.1. Identification of novel binding sites

Both the crystal structure of the immature complex (PDB file: 3N40[95]) and the mature complex (PDB file: 3N42[95]) were used. Binding sites within the receptors were detected using the Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012). The algorithm is based on a grid search and "eraser" algorithm which derives binding sites from cavities in the structure of the receptor. The binding site found is displayed as a set of points. The volume of each cavity is defined as the product of

number of site points and the cube of the grid spacing. Six main sites were detected in both the immature and the mature crystal structures and only one site were detected in the mature crystal structure that is not present in the immature form (Figure 6.2). Table 1 shows the identified sites with their characters. Suitable cavities were then checked further based on functionality, presence of hydrophobic residues, presence of charged residues and solvent accessibility.

9.2.2.2. Virtual screening with the CHIKV envelope proteins

Two chemical compounds libraries were used; The NCI set library of 265,242 compounds and the Life chemicals protein-protein interactions inhibitors library of 31,143 compounds. The databases were filtered with the drug-likeness-index; limit the range for Molecular Weight $\leq$ 500, calculated octanol–water partition coefficient (clogP $\leq$ 5), and hydrogen bond donors, and acceptors (OH's and NH's $\leq$ 5; N's and O's $\leq$ 10),[231] using Filter v2.0.2 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com), producing 55,841 compounds from the NCI library and 4,124 compounds from the Life Chemicals library. Fast exhaustive virtual screening was performed using FRED v2.2.5 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com). FRED is a fast and effective docking application whose performance is significantly more reliable, i.e. lower variance, than most other programs.[232, 233] FRED performs a systematic, exhaustive, nonstochastic examination of all possible poses within the protein active site, filters for shape complementarily [234] and pharmacophoric features before selecting and optimizing poses using the Chemgauss scoring function. Omega2 (Systematic high-throughput conformer generation, OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com),[235, 236] was used to generate multiple conformers for each compound in the database libraries using the default settings. Omega2 takes into account the flexibility of a molecule by generating all representative conformers. For the NCI library, 2,312,012 conformers were generated, and 334,064 conformers were generated from the Life Chemicals compounds. The work-flow diagram is shown in (Figure 6.3), the life chemical library was screened on site 2 (light green, Figure 6.2) in both of the immature and the mature glycoproteins. The NCI set compounds were screened on site 4 (blue, Figure 6.2) of the two envelope protein forms. The binding sites were prepared for docking using Fred receptor setup software (OpenEye Scientific Software, Santa Fe,

NM. http://www.eyesopen.com). The grid boxes were determined based on the x, y and z co-ordinates given in Table 1. For site 2 in the 3N40 receptor, the box size was set to 6153 $Å^3$ and was assigned an inner contour of 99 $Å^3$ and an outer contour of 1886 $Å^3$. Site 4 in the 3N40 receptor has a box size of 6580 $Å^3$ and was assigned an inner contour of 116 $Å^3$ and an outer contour of 1071 $Å^3$. Site 2 in 3N42 receptor has a box size of 7578 $Å^3$ and was assigned an inner contour of 66 $Å^3$ and an outer contour of 1816 $Å^3$. Site 4 in 3N42 receptor was assigned a box size of 6482 $Å^3$, an inner contour of 45 $Å^3$ and an outer contour of 1547 $Å^3$. No constraints were enabled in any of the prepared receptors. During the docking calculations, both chemgauss3 and shapeguass scoring functions were enabled. After the docking calculations, the poses returned were scored and ranked with a Gaussian shape function independently by the five available scoring functions (PLP, Chemgauss3, Chemscore, OEChemscore, and Screenscore) and by a consensus of all. The top ranked poses from the exhaustive docking were then optimized using systematic solid body optimization by chemgauss3. VIDA v4.2.0 (OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com) was used to visualise the docked poses within the receptor active site, and to inspect the critical interacting residues in each pocket with the individual docked poses. Top 20 hits were then recorded for each of the four sites **(**Tables 6.2, 6.5, 6.8, 6.11**).**

The top 20 docked poses ranked in each of the four binding sites were then extracted as PDB files, and were processed with AutoDock Tools 1.5.6rc3 (ADT) graphical interface.[173] The Gasteiger charges were computed and the nonpolar hydrogen atoms were merged, torsion angles were defined, they were then saved as pdbqt files for Autodock calculations. Crystal structures (3N40, 3N42) were used by AutoDock Tools 1.5.6rc3 to setup the receptor binding sites. The grid box co-ordinates in each site were determined based on the co-ordinates in Table 6.1. The grid box size was set to 46 x 46 x 46 points in x, y, and a z direction in each of the four sites and a grid spacing of 0.375 Å was used. AutoGrid 4.2 algorithm was used to evaluate the binding energies between the inhibitors and the enzyme and to generate the energy maps for the docking run. Fifty runs were generated by using Autodock 4.2 Lamarckian genetic algorithm[173] for the searches. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 2.0 Å, the docked poses were ranked according to the binding energies and ligand efficiencies, and finally the five lowest energy poses **(**Tables 6.3, 6.4, 6.6, 6.7, 6.9, 6.10, 6.12, 6.13**)** were selected as the resultant complexes with the enzymes. The

complexes were then typed with the CHARMm forcefield with Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012) to relax the obtained poses within the enzyme pockets, and visualized.

9.2.3. CHIKV non-structural protein 3 (nsP3)

9.2.3.1. Virtual screening with Autodock Vina

The CHIKV nsP3 macro domain crystal structure was downloaded from the protein data bank (RCSB), PDB code: 3GPO.[80] Autodock Tools 1.5.6rc3 (ADT) graphical interface[173] was used to prepare the receptor, where waters were removed, non-polar hydrogens were merged and polar hydrogens were added. The Kollman charges were added and the Gasteiger charges were calculated for the receptor. The grid box co-ordinates were determined according to the co-crystallized ligand (ADPribose) where x, y and z values were 29.642, 29.232 and 21.592, respectively. The grid box size was set to 32 x 32 x 32 Å, with a grid spacing of 0.375 Å. The receptor was then saved as a pdbqt file, and was ready for the virtual screening run.

The NCI Diversity Set III 1990 compounds' library was downloaded from the NCI website (http://cactus.nci.nih.gov/download/nci/). Compounds were energy minimized using the Uff forcefield[267] embedded within the Open Babel of PyRx (Scripps research institute, http://pyrx.sourceforge.net/), with 100,000 steps where the minimization stopped when the energy difference was less than 0.001. The minimized structures were then converted to the pdbqt format using Autodock Tools 1.5.6rc3 (ADT).[173]

Autodock Vina was used to dock each compound in the library within the nsP3 binding site, using an exhaustiveness (number of runs) value of 125. The docked poses were then ranked according to the Vina binding energies (Kcal/mol). Top poses ranked that achieved better binding energies than the co-crystallized ligand (ADP-ribose), were extracted, complexed with the receptor (nsP3). The complexes were then typed with the CHARMm forcefield with Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012) to relax the obtained poses within the enzyme pocket, and visualized.

9.2.3.2. Re-ranking using Autodock 4

The Autodock Vina hit list poses (Table 3.1) were extracted as pdbqt files by (ADT). The CHIKV nsP3 receptor that was prepared for Autodock Vina was re-used for Autodock 4. AutoGrid 4.2 algorithm was used to evaluate the binding energies between the inhibitors and the enzyme and to generate the energy maps for the docking run. Fifty runs were generated by using Autodock 4.2 Lamarckian genetic algorithm[173] for the searches. Cluster analysis was performed on docked results, with a root-mean-square tolerance of 1.0, 2.0 and 3.0 Å, the docked poses were ranked according to the binding energies and ligand efficiencies, and finally, the survived lowest energy poses (Table 3.2) were selected as the resultant complexes with the enzyme. The complexes were then typed with the CHARMm forcefield with Discovery Studio 3.5 software (Accelrys Software Inc.: San Diego, CA, 2012) to relax the obtained poses within the enzyme pocket, and visualized.

9.3. Biological evaluation

9.3.1. Anti-Chikungunya evaluations

Currently, the synthetic compounds **114**-**117**, **120**, **121**, **126**, **127**, **130**, **131**, **137**, **138**, **140**-**143** and **145**-**147**, along with the CHIKV nsP2 data base searching hits (Tables 5.3 and 5.6) are being evaluated for their anti-CHIKV activity including virus-cell-based CPE reduction and cytotoxicity assays (with the standard deviations). Testing is being conducted at Rega Institute for Medical Research, University of Leuven (KU Leuven), Belgium, with Prof. Johan Neyts and Dr. Peiter Leyssen.

The NCI hits (Table 3.1) are being evaluated against the CHIKV by Professor Suresh Mahalingam, Emerging Viruses and Inflammation Research Group, Institute for Glycomics, Griffith University, Gold Coast, QLD, Australia, 4222.

9.3.2. Trypanocidal activity evaluation

9.3.2.1. Initial screen

The bis-sulfonamide compounds **160**, **170**-**194** were resuspended in the appropriate volume of 100% DMSO to give a stock solution of 21 mM. Initially all compounds

were screened at doses of 10.43 µM and 1.4 µM against *T. b. brucei*, to give an indication of the compounds trypanocidal activity. The compounds were screened once against two separate trypanosome populations, designated culture A and culture B. Compounds with >80% activity at 10.43 µM and >50% activity at 1.4 µM were classified as active. Using these selection criteria, active compounds were selected for IC_{50} and selectivity index (SI) determination.

9.3.2.2. Assay controls

Positive in-plate controls in wells G2-O23 contained 0.417% final DMSO concentration. The in-plate controls were used to calculate the activity of compounds whilst taking into account any plate-to-plate variability in the signal of the assay. Two types of external control plates were included for each trypanosome population. To calculate the assay Z', a plate containing minimum and maximum assay signals was used. For calculation of the minimum signal, or 100% inhibition of cell growth, 36.45 µM final concentration puromycin was placed in half of the plate. To calculate the maximum signal, 0.42% DMSO was placed in the second half of the plate. The Z' prime was 0.57 and 0.72 for culture A and B respectively.

As a measure of the sensitivity of each assay, a plate containing dose response dilutions of the reference compounds pentamidine, diminazene aceturate and puromycin was used. Puromycin gave IC_{50} values of 54.78 nM and 73.69 nM for culture A and B, respectively. Pentamidine exhibited an IC_{50} value of 1.6 nM and 2.44 nM for culture A and B respectively. Diminazene aceturate had an IC_{50} value of 103.8 nM and 155.2 nM.

9.3.2.3. IC_{50} and selectivity index determination

The compounds were screened against *T. b. brucei* and HEK 293 in 20 point dose CRC format. The highest dose of 83.33 µM was only used in the HEK293 assay. This allowed the selectivity index of the compounds to be estimated. The compounds were screened in duplicate against two separate cultures (designated A and B) of both *T. brucei* and HEK 293. The IC_{50} and SI was determined for compounds with greater than >80% activity at 41.67 µM in the *T. b. brucei* assay. The SI of the compound was determined where possible by directly comparing the IC_{50} values between the two assays. If this was not possible, an estimated IC_{50} value was calculated by comparing

the dose at which the compound was active >50% in the *T. brucei* assay and the lowest dose at which there was no activity (<50%) in the HEK293 assay.

Assay controls: Positive in-plate controls in column 23 (rows B-O) contained 0.417% final DMSO concentration. The in-plate controls were used to calculate the activity of compounds whilst taking into account any plate-to-plate variability in the signal of the assay. Two types of external control plates were included for each of the two cell populations in the *T. b. brucei* and HEK 293 assay. To calculate the assay Z', a plate containing minimum and maximum assay signals was used. For calculation of the minimum signal, or 100% inhibition of cell growth, 36.45 μM final puromycin was placed in half of the plate. To calculate the maximum signal, 0.42% DMSO was placed in the second half of the plate. The Z' prime was 0.63 and 0.72 for culture A and B in the T. brucei assay respectively. In the HEK 293 assay the Z' prime was 0.69 for culture A and 0.56 for culture B. As a measure of the sensitivity of each assay, a plate containing dose response dilutions of the reference compounds pentamidine, diminazene aceturate and puromycin was used. In the *T. brucei* assay Puromycin gave IC_{50} values of 43.6 nM and 47.62 nM for culture A and B, respectively. Pentamidine exhibited an IC_{50} value of 1.58 nM and 1.78 nM for culture A and B respectively. Diminazene aceturate had an IC_{50} value of 94.6 nM and 88.9 nM. In the HEK 293 assay, Puromycin gave IC_{50} values of 165.3 nM and 225.7 nM for culture A and B, respectively. Pentamidine and diminazene aceturate exhibited no activity in the HEK 293 assay.

CHAPTER 10: References

1. Her, Z.; Kam, Y. W.; Lin, R. T.; Ng, L. F. Chikungunya: a bending reality. *Microb. Infect.* **2009**, 11, 1165-1176.
2. Robinson, M. C. An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952-53. I. Clinical features. *Trans. R. Soc. Trop. Med. Hyg.* **1955**, 49, 28-32.
3. Lumsden, W. H. An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952-53. II. General description and epidemiology. *Trans. R. Soc. Trop. Med. Hyg.* **1955**, 49, 33-57.
4. Nimmannitya, S.; Halstead, S. B.; Cohen, S. N.; Margiotta, M. R. Dengue and chikungunya virus infection in man in Thailand, 1962-1964. I. Observations on hospitalized patients with hemorrhagic fever. *Am. J. Trop. Med. Hyg.* **1969**, 18, 954-971.
5. Njenga, M. K.; Nderitu, L.; Ledermann, J. P.; Ndirangu, A.; Logue, C. H.; Kelly, C. H. L.; Sang, R.; Sergon, K.; Breiman, R.; Powers, A. M. Tracking epidemic Chikungunya virus into the Indian Ocean from East Africa. *J. Gen. Virol.* **2008**, 89, 2754-2760.
6. Ravi, V. Re-emergence of chikungunya virus in India. *Indian J. Med. Microbiol.* **2006**, 24, 83-84.
7. Sambri, V.; Cavrini, F.; Rossini, G.; Pierro, A.; Landini, M. P. The 2007 epidemic outbreak of Chikungunya virus infection in the Romagna region of Italy: a new perspective for the possible diffusion of tropical diseases in temperate areas? *New Microbiol.* **2008**, 31, 303-304.
8. Enserink, M. Infectious diseases - Chikungunya: No longer a third world disease. *Science* **2007**, 318, 1860-1861.
9. Powers, A. M.; Logue, C. H. Changing patterns of chikungunya virus: re-emergence of a zoonotic arbovirus. *J. Gen. Virol.* **2007**, 88, 2363-2377.
10. Schwartz, O.; Albert, M. L. Biology and pathogenesis of chikungunya virus. *Nat. Rev. Microbiol.* **2010**, 8, 491-500.
11. Outbreak news. Chikungunya, India. *Wkly. Epidemiol. Rec.* **2006**, 81, 409-410.

12. Saxena, S. K.; Singh, M.; Mishra, N.; Lakshmi, V. Resurgence of chikungunya virus in India: an emerging threat. *Euro Surveill.* **2006**, 11(8), E060810.2.
13. Chhabra, M.; Mittal, V.; Bhattacharya, D.; Rana, U.; Lal, S. Chikungunya fever: a re-emerging viral infection. *Indian J. Med. Microbiol.* **2008**, 26, 5-12.
14. Vazeille, M.; Moutailler, S.; Coudrier, D.; Rousseaux, C.; Khun, H.; Huerre, M.; Thiria, J.; Dehecq, J. S.; Fontenille, D.; Schuffenecker, I.; Despres, P.; Failloux, A. B. Two chikungunya isolates from the outbreak of La Reunion (Indian Ocean) exhibit different patterns of infection in the mosquito, Aedes albopictus. *PLoS One* **2007**, 2, e1168.
15. Pages, F.; Peyrefitte, C. N.; Mve, M. T.; Jarjaval, F.; Brisse, S.; Iteman, I.; Gravier, P.; Tolou, H.; Nkoghe, D.; Grandadam, M. Aedes albopictus mosquito: the main vector of the 2007 chikungunya outbreak in Gabon. *PLoS One* **2009**, 4, e4691.
16. Gratz, N. G. Critical review of the vector status of Aedes albopictus. *Med. Vet. Entomol.* **2004**, 18, 215-227.
17. Smith, C. E. The history of dengue in tropical Asia and its probable relationship to the mosquito Aedes aegypti. *J. Trop. Med. Hyg.* **1956**, 59, 243-251.
18. Gerardin, P.; Barau, G.; Michault, A.; Bintner, M.; Randrianaivo, H.; Choker, G.; Lenglet, Y.; Touret, Y.; Bouveret, A.; Grivard, P.; Le Roux, K.; Blanc, S.; Schuffenecker, I.; Couderc, T.; Arenzana-Seisdedos, F.; Lecuit, M.; Robillard, P. Y. Multidisciplinary prospective study of mother-to-child chikungunya virus infections on the island of La Reunion. *PLoS Med.* **2008**, 5, e60.
19. Ziegler, S. A.; Lu, L.; da Rosa, A. P.; Xiao, S. Y.; Tesh, R. B. An animal model for studying the pathogenesis of chikungunya virus infection. *Am. J. Trop. Med. Hyg.* **2008**, 79, 133-139.
20. Queyriaux, B.; Simon, F.; Grandadam, M.; Michel, R.; Tolou, H.; Boutin, J. P. Clinical burden of chikungunya virus infection. *Lancet Infect. Dis.* **2008**, 8, 2-3.
21. Santhosh, S. R.; Dash, P. K.; Parida, M. M.; Khan, M.; Tiwari, M.; Lakshmana Rao, P. V. Comparative full genome analysis revealed E1: A226V shift in 2007 Indian Chikungunya virus isolates. *Virus Res.* **2008**, 135, 36-41.
22. Robin, S.; Ramful, D.; Le Seach, F.; Jaffar-Bandjee, M. C.; Rigou, G.; Alessandri, J. L. Neurologic manifestations of pediatric chikungunya infection. *J. Child Neurol.* **2008**, 23, 1028-1035.

23. Chandak, N. H.; Kashyap, R. S.; Kabra, D.; Karandikar, P.; Saha, S. S.; Morey, S. H.; Purohit, H. J.; Taori, G. M.; Daginawala, H. F. Neurological complications of chikungunya virus infection. *Neurol. India* **2009**, 57, 177-180.
24. Pialoux, G.; Gauzere, B. A.; Jaureguiberry, S.; Strobel, M. Chikungunya, an epidemic arbovirosis. *Lancet Infect. Dis.* **2007**, 7, 319-327.
25. Mahesh, G.; Giridhar, A.; Shedbele, A.; Kumar, R.; Saikumar, S. J. A case of bilateral presumed chikungunya neuroretinitis. *Indian J. Ophthalmol.* **2009**, 57, 148-150.
26. Nair, A. G.; Biswas, J.; Bhende, M. P. A case of bilateral chikungunya neuroretinitis. *J. Ophthalmic. Inflamm. Infect.* **2012**, 2, 39-40.
27. Couderc, T.; Gangneux, N.; Chretien, F.; Caro, V.; Le Luong, T.; Ducloux, B.; Tolou, H.; Lecuit, M.; Grandadam, M. Chikungunya virus infection of corneal grafts. *J. Infect. Dis.* **2012**, 206, 851-859.
28. Singh, S. K.; Unni, S. K. Chikungunya virus: host pathogen interaction. *Rev. Med. Virol.* **2011**, 21, 78-88.
29. Grakoui, A.; Levis, R.; Raju, R.; Huang, H. V.; Rice, C. M. A Cis-Acting Mutation in the Sindbis Virus Junction Region Which Affects subgenomic RNA-Synthesis. *J. Virol.* **1989**, 63, 5216-5227.
30. Sourisseau, M.; Schilte, C.; Casartelli, N.; Trouillet, C.; Guivel-Benhassine, F.; Rudnicka, D.; Sol-Foulon, N.; Le Roux, K.; Prevost, M. C.; Fsihi, H.; Frenkiel, M. P.; Blanchet, F.; Afonso, P. V.; Ceccaldi, P. E.; Ozden, S.; Gessain, A.; Schuffenecker, I.; Verhasselt, B.; Zamborlini, A.; Saib, A.; Rey, F. A.; Arenzana-Seisdedos, F.; Despres, P.; Michault, A.; Albert, M. L.; Schwartz, O. Characterization of reemerging chikungunya virus. *PLoS Path.* **2007**, 3, e89.
31. Wintachai, P.; Wikan, N.; Kuadkitkan, A.; Jaimipuk, T.; Ubol, S.; Pulmanausahakul, R.; Auewarakul, P.; Kasinrerk, W.; Weng, W. Y.; Panyasrivanit, M.; Paemanee, A.; Kittisenachai, S.; Roytrakul, S.; Smith, D. R. Identification of prohibitin as a Chikungunya virus receptor protein. *J. Med. Virol.* **2012**, 84, 1757-1770.
32. Berger, K. H.; Yaffe, M. P. Prohibitin family members interact genetically with mitochondrial inheritance components in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **1998**, 18, 4043-4052.

33. Kolonin, M. G.; Saha, P. K.; Chan, L.; Pasqualini, R.; Arap, W. Reversal of obesity by targeted ablation of adipose tissue. *Nat. Med.* **2004**, 10, 625-632.
34. Kielian, M.; Rey, F. A. Virus membrane-fusion proteins: more than one way to make a hairpin. *Nat. Rev. Microbiol.* **2006**, 4, 67-76.
35. Marsh, M.; Helenius, A. Virus entry: open sesame. *Cell* **2006**, 124, 729-740.
36. Chatterjee, P. K.; Vashishtha, M.; Kielian, M. Biochemical consequences of a mutation that controls the cholesterol dependence of Semliki Forest virus fusion. *J. Virol.* **2000**, 74, 1623-1631.
37. Smit, J. M.; Bittman, R.; Wilschut, J. Low-pH-dependent fusion of Sindbis virus with receptor-free cholesterol- and sphingolipid-containing liposomes. *J. Virol.* **1999**, 73, 8476-8484.
38. Barton, D. J.; Sawicki, S. G.; Sawicki, D. L. Solubilization and immunoprecipitation of alphavirus replication complexes. *J. Virol.* **1991**, 65, 1496-1506.
39. Shirako, Y.; Strauss, J. H. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J. Virol.* **1994**, 68, 1874-1885.
40. Schilte, C.; Couderc, T.; Chretien, F.; Sourisseau, M.; Gangneux, N.; Guivel-Benhassine, F.; Kraxner, A.; Tschopp, J.; Higgs, S.; Michault, A.; Arenzana-Seisdedos, F.; Colonna, M.; Peduto, L.; Schwartz, O.; Lecuit, M.; Albert, M. L. Type I IFN controls chikungunya virus via its action on nonhematopoietic cells. *J. Exp. Med.* **2010**, 207, 429-442.
41. Khan, A. H.; Morita, K.; Parquet Md Mdel, C.; Hasebe, F.; Mathenge, E. G.; Igarashi, A. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *J. Gen. Virol.* **2002**, 83, 3075-3084.
42. Tang, B. L. The cell biology of chikungunya virus infection. *Cell. Microbiol.* **2012**, 14, 1354-1363.
43. Perera, R.; Owen, K. E.; Tellinghuisen, T. L.; Gorbalenya, A. E.; Kuhn, R. J. Alphavirus nucleocapsid protein contains a putative coiled coil alpha-helix important for core assembly. *J. Virol.* **2001**, 75, 1-10.
44. Maek, A. N. W.; Silachamroon, U. Presence of autoimmune antibody in chikungunya infection. *Case Report Med.* **2009**, 2009, 840183.

45. Eckels, K. H.; Harrison, V. R.; Hetrick, F. M. Chikungunya virus vaccine prepared by Tween-ether extraction. *Appl. Microbiol.* **1970**, 19, 321-325.
46. Levitt, N. H.; Ramsburg, H. H.; Hasty, S. E.; Repik, P. M.; Cole, F. E.; Lupton, H. W. Development of an attenuated strain of chikungunya virus for use in vaccine production. *Vaccine* **1986**, 4, 157-162.
47. McClain, D. J.; Pittman, P. R.; Ramsburg, H. H.; Nelson, G. O.; Rossi, C. A.; Mangiafico, J. A.; Schmaljohn, A. L.; Malinoski, F. J. Immunologic interference from sequential administration of live attenuated alphavirus vaccines. *J. Infect. Dis.* **1998**, 177, 634-641.
48. Edelman, R.; Tacket, C. O.; Wasserman, S. S.; Bodison, S. A.; Perry, J. G.; Mangiafico, J. A. Phase II safety and immunogenicity study of live chikungunya virus vaccine TSI-GSD-218. *Am. J. Trop. Med. Hyg.* **2000**, 62, 681-685.
49. Tiwari, M.; Parida, M.; Santhosh, S. R.; Khan, M.; Dash, P. K.; Rao, P. V. Assessment of immunogenic potential of vero adapted formalin inactivated vaccine derived from novel ECSA genotype of chikungunya virus. *Vaccine* **2009**, 27, 2513-2522.
50. Plante, K.; Wang, E. Y.; Partidos, C. D.; Weger, J.; Gorchakov, R.; Tsetsarkin, K.; Borland, E. M.; Powers, A. M.; Seymour, R.; Stinchcomb, D. T.; Osorio, J. E.; Frolov, I.; Weaver, S. C. Novel chikungunya vaccine candidate with an IRES-based attenuation and host range alteration mechanism. *PLoS Path.* **2011**, 7, e1002142.
51. Velez, R. A.; de Matos, A. P. A.; Parreira, R.; Piedade, J.; Matos, B.; Correia, C.; Esteves, A. Expression of chikungunya virus-like particles. *Microsc. Microanal.* **2012**, 18, 59-60.
52. Akahata, W.; Nabel, G. J. A Specific domain of the chikungunya virus E2 protein regulates particle formation in human cells: Implications for alphavirus vaccine design. *J. Virol.* **2012**, 86, 8879-8883.
53. Akahata, W.; Yang, Z. Y.; Andersen, H.; Sun, S.; Holdaway, H. A.; Kong, W. P.; Lewis, M. G.; Higgs, S.; Rossmann, M. G.; Rao, S.; Nabel, G. J. A virus-like particle vaccine for epidemic chikungunya virus protects nonhuman primates against infection. *Nat. Med.* **2010**, 16, 334-338.
54. Solignat, M.; Gay, B.; Higgs, S.; Briant, L.; Devaux, C. Replication cycle of chikungunya: a re-emerging arbovirus. *Virology* **2009**, 393, 183-197.

55. Jones, P. H.; Maric, M.; Madison, M. N.; Maury, W.; Roller, R. J.; Okeoma, C. M. BST-2/tetherin-mediated restriction of chikungunya (CHIKV) VLP budding is counteracted by CHIKV non-structural protein 1 (nsP1). *Virology* **2013**, 438, 37-49.
56. Jones, P. H.; Mehta, H. V.; Maric, M.; Roller, R. J.; Okeoma, C. M. Bone marrow stromal cell antigen 2 (BST-2) restricts mouse mammary tumor virus (MMTV) replication in vivo. *Retrovirology* **2012**, 9:10.
57. Hardy, W. R.; Strauss, J. H. Processing the nonstructural polyproteins of Sindbis virus - nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J. Virol.* **1989**, 63, 4653-4664.
58. Merits, A.; Vasiljeva, L.; Ahola, T.; Kaariainen, L.; Auvinen, P. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *J. Gen. Virol.* **2001**, 82, 765-773.
59. Strauss, E. G.; De Groot, R. J.; Levinson, R.; Strauss, J. H. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* **1992**, 191, 932-940.
60. Vasiljeva, L.; Valmu, L.; Kaariainen, L.; Merits, A. Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *J. Biol. Chem.* **2001**, 276, 30786-30793.
61. Gorbalenya, A. E.; Koonin, E. V. Helicases - Amino-acid-sequence comparisons and structure-function-relationships. *Curr. Opin. Struct. Biol.* **1993**, 3, 419-429.
62. Kim, K. H.; Rumenapf, T.; Strauss, E. G.; Strauss, J. H. Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology* **2004**, 323, 153-163.
63. Vasiljeva, L.; Merits, A.; Golubtsov, A.; Sizemskaja, V.; Kaariainen, L.; Ahola, T. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J. Biol. Chem.* **2003**, 278, 41636-41645.
64. Pastorino, B. A.; Peyrefitte, C. N.; Almeras, L.; Grandadam, M.; Rolland, D.; Tolou, H. J.; Bessaud, M. Expression and biochemical characterization of nsP2 cysteine protease of chikungunya virus. *Virus Res.* **2008**, 131, 293-298.
65. Karpe, Y. A.; Aher, P. P.; Lole, K. S. NTPase and 5 '-RNA triphosphatase activities of chikungunya virus nsP2 protein. *PLoS One* **2011**, 6, e22336.

66. Domsalla, A.; Melzig, M. F. Occurrence and properties of proteases in plant latices. *Planta Med.* **2008**, 74, 699-711.
67. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, 28, 235-242.
68. Grudkowska, M.; Zagdanska, B. Multifunctional role of plant cysteine proteinases. *Acta Biochim. Pol.* **2004**, 51, 609-624.
69. Singh Kh, D.; Kirubakaran, P.; Nagarajan, S.; Sakkiah, S.; Muthusamy, K.; Velmurgan, D.; Jeyakanthan, J. Homology modeling, molecular dynamics, e-pharmacophore mapping and docking study of chikungunya virus nsP2 protease. *J. Mol. Model.* **2012**, 18, 39-51.
70. Russo, A. T.; Malmstrom, R. D.; White, M. A.; Watowich, S. J. Structural basis for substrate specificity of alphavirus nsP2 proteases. *J. Mol. Graphics Model.* **2010**, 29, 46-53.
71. Bassetto, M.; De Burghgraeve, T.; Delang, L.; Massarotti, A.; Coluccia, A.; Zonta, N.; Gatti, V.; Colombano, G.; Sorba, G.; Silvestri, R.; Tron, G. C.; Neyts, J.; Leyssen, P.; Brancale, A. Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. *Antiviral Res.* **2013**, 98, 12-18.
72. Breakwell, L.; Dosenovic, P.; Karlsson Hedestam, G. B.; D'Amato, M.; Liljestrom, P.; Fazakerley, J.; McInerney, G. M. Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J. Virol.* **2007**, 81, 8677-8684.
73. Frolov, I.; Garmashova, N.; Atasheva, S.; Frolova, E. I. Random Insertion Mutagenesis of Sindbis Virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. *J. Virol.* **2009**, 83, 9031-9044.
74. Bourai, M.; Lucas-Hourani, M.; Gad, H. H.; Drosten, C.; Jacob, Y.; Tafforeau, L.; Cassonnet, P.; Jones, L. M.; Judith, D.; Couderc, T.; Lecuit, M.; Andre, P.; Kummerer, B. M.; Lotteau, V.; Despres, P.; Tangy, F.; Vidalain, P. O. Mapping of chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. *J. Virol.* **2012**, 86, 3121-3134.
75. De, I.; Fata-Hartley, C.; Sawicki, S. G.; Sawicki, D. L. Functional analysis of nsP3 phosphoprotein mutants of Sindbis virus. *J. Virol.* **2003**, 77, 13106-13116.

76. Lastarza, M. W.; Grakoui, A.; Rice, C. M. Deletion and duplication mutations in the C-terminal nonconserved region of Sindbis virus nsP3: effects on phosphorylation and on virus replication in vertebrate and invertebrate cells. *Virology* **1994**, 202, 224-232.
77. Li, G. P.; La Starza, M. W.; Hardy, W. R.; Strauss, J. H.; Rice, C. M. Phosphorylation of Sindbis virus nsP3 in vivo and in vitro. *Virology* **1990**, 179, 416-427.
78. Vihinen, H.; Ahola, T.; Tuittila, M.; Merits, A.; Kaariainen, L. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. *J. Biol. Chem.* **2001**, 276, 5745-5752.
79. Tuittila, M. T.; Santagati, M. G.; Roytta, M.; Maatta, J. A.; Hinkkanen, A. E. Replicase complex genes of Semliki Forest virus confer lethal neurovirulence. *J. Virol.* **2000**, 74, 4579-4589.
80. Malet, H.; Coutard, B.; Jamal, S.; Dutartre, H.; Papageorgiou, N.; Neuvonen, M.; Ahola, T.; Forrester, N.; Gould, E. A.; Lafitte, D.; Ferron, F.; Lescar, J.; Gorbalenya, A. E.; de Lamballerie, X.; Canard, B. The crystal structures of chikungunya and venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *J. Virol.* **2009**, 83, 6534-6545.
81. Neuvonen, M.; Ahola, T. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. *J. Mol. Biol.* **2009**, 385, 212-225.
82. Lulla, A.; Lulla, V.; Merits, A. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J. Virol.* **2012**, 86, 553-565.
83. Shin, G.; Yost, S. A.; Miller, M. T.; Elrod, E. J.; Grakoui, A.; Marcotrigiano, J. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proc. Natl. Acad. Sci. USA* **2012**, 109, 16534-16539.
84. Vasiljeva, L.; Merits, A.; Golubtsov, A.; Sizemskaja, V.; Kaariainen, L.; Ahola, T. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J. Biol. Chem.* **2003**, 278, 41636-41645.
85. Rungrotmongkol, T.; Nunthaboot, N.; Malaisree, M.; Kaiyawet, N.; Yotmanee, P.; Meeprasert, A.; Hannongbua, S. Molecular insight into the specific binding of ADP-ribose to the nsP3 macro domains of chikungunya and venezuelan equine

encephalitis viruses: Molecular dynamics simulations and free energy calculations. *J. Mol. Graphics Model.* **2010**, 29, 347-353.

86. Neuvonen, M.; Kazlauskas, A.; Martikainen, M.; Hinkkanen, A.; Ahola, T.; Saksela, K. SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. *PLoS Path.* **2011**, 7, e1002383.
87. Fros, J. J.; Domeradzka, N. E.; Baggen, J.; Geertsema, C.; Flipse, J.; Vlak, J. M.; Pijlman, G. P. Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci. *J. Virol.* **2012**, 86, 10873-10879.
88. Parker, F.; Maurier, F.; Delumeau, I.; Duchesne, M.; Faucher, D.; Debussche, L.; Dugue, A.; Schweighoffer, F.; Tocque, B. A Ras-GTPase-activating protein SH3-domain-binding protein. *Mol. Cell. Biol.* **1996**, 16, 2561-2569.
89. Anderson, P.; Kedersha, N. Stress granules. *Curr. Biol.* **2009**, 19, R397-398.
90. Shirako, Y.; Strauss, E. G.; Strauss, J. H. Suppressor mutations that allow Sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: Evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. *Virology* **2000**, 276, 148-160.
91. Rathore, A. P.; Ng, M. L.; Vasudevan, S. G. Differential unfolded protein response during chikungunya and Sindbis virus infection: CHIKV nsP4 suppresses eIF2alpha phosphorylation. *Virol. J.* **2013**, 10:36.
92. Tardif, K. D.; Waris, G.; Siddiqui, A. Hepatitis C virus, ER stress, and oxidative stress. *Trends Microbiol.* **2005**, 13, 159-163.
93. Salminen, A.; Wahlberg, J. M.; Lobigs, M.; Liljestrom, P.; Garoff, H. Membrane fusion process of Semliki Forest virus. II: Cleavage-dependent reorganization of the spike protein complex controls virus entry. *J. Cell. Biol.* **1992**, 116, 349-357.
94. Li, L.; Jose, J.; Xiang, Y.; Kuhn, R. J.; Rossmann, M. G. Structural changes of envelope proteins during alphavirus fusion. *Nature* **2010**, 468, 705-708.
95. Voss, J. E.; Vaney, M. C.; Duquerroy, S.; Vonrhein, C.; Girard-Blanc, C.; Crublet, E.; Thompson, A.; Bricogne, G.; Rey, F. A. Glycoprotein organization of chikungunya virus particles revealed by X-ray crystallography. *Nature* **2010**, 468, 709-712.
96. Kuo, S. C.; Chen, Y. J.; Wang, Y. M.; Tsui, P. Y.; Kuo, M. D.; Wu, T. Y.; Lo, S. J. Cell-based analysis of chikungunya virus E1 protein in membrane fusion. *J. Biomed. Sci.* **2012**, 19:44.

97. Kielian, M.; Rey, F. A. Virus membrane-fusion proteins: more than one way to make a hairpin. *Nat. Rev. Microbiol.* **2006**, 4, 67-76.

98. Weissenhorn, W.; Hinz, A.; Gaudin, Y. Virus membrane fusion. *FEBS Lett.* **2007**, 581, 2150-2155.

99. Mohanram, H.; Nip, A.; Domadia, P. N.; Bhunia, A.; Bhattacharjya, S. NMR structure, localization, and vesicle fusion of chikungunya virus fusion peptide. *Biochemistry* **2012**, 51, 7863-7872.

100. Ozden, S.; Lucas-Hourani, M.; Ceccaldi, P. E.; Basak, A.; Valentine, M.; Benjannet, S.; Hamelin, J.; Jacob, Y.; Mamchaoui, K.; Mouly, V.; Despres, P.; Gessain, A.; Butler-Browne, G.; Chretien, M.; Tangy, F.; Vidalain, P. O.; Seidah, N. G. Inhibition of chikungunya virus infection in cultured human muscle cells by furin inhibitors: impairment of the maturation of the E2 surface glycoprotein. *J. Biol. Chem.* **2008**, 283, 21899-21908.

101. Lobigs, M.; Zhao, H. X.; Garoff, H. Function of Semliki Forest virus E3 peptide in virus assembly: replacement of E3 with an artificial signal peptide abolishes spike heterodimerization and surface expression of E1. *J. Virol.* **1990**, 64, 4346-4355.

102. Meyer, W. J.; Johnston, R. E. Structural Rearrangement of Infecting Sindbis Virions at the Cell-Surface - Mapping of newly accessible epitopes. *J. Virol.* **1993**, 67, 5117-5125.

103. Meyer, W. J.; Gidwitz, S.; Ayers, V. K.; Schoepp, R. J.; Johnston, R. E. Conformational alteration of Sindbis virion glycoproteins induced by heat, reducing agents, or low pH. *J. Virol.* **1992**, 66, 3504-3513.

104. Flynn, D. C.; Meyer, W. J.; Mackenzie, J. M., Jr.; Johnston, R. E. A conformational change in Sindbis virus glycoproteins E1 and E2 is detected at the plasma membrane as a consequence of early virus-cell interaction. *J. Virol.* **1990**, 64, 3643-3653.

105. de Lamballerie, X.; Ninove, L.; Charrel, R. N. Antiviral treatment of chikungunya virus infection. *Infect. Disord. Drug Targets* **2009**, 9, 101-104.

106. Couderc, T.; Chretien, F.; Schilte, C.; Disson, O.; Brigitte, M.; Guivel-Benhassine, F.; Touret, Y.; Barau, G.; Cayet, N.; Schuffenecker, I.; Despres, P.; Arenzana-Seisdedos, F.; Michault, A.; Albert, M. L.; Lecuit, M. A mouse model

for chikungunya: young age and inefficient type-I interferon signaling are risk factors for severe disease. *PLoS Path.* **2008**, 4, e29.

107. Yeo, L. S.; Chu, J. J. H. Recent developments and challenges in mouse models of Chikungunya virus infection. *Future Virol.* **2013**, 8, 423-426.
108. Inglot, A. D. Comparison of the antiviral activity in vitro of some non-steroidal anti-inflammatory drugs. *J. Gen. Virol.* **1968**, 4, 203-214.
109. Shimizu, Y.; Yamamoto, S.; Homma, M.; Ishida, N. Effect of chloroquine on the growth of animal viruses. *Arch Gesamte Virusforsch.* **1972**, 36, 93-104.
110. Coombs, K.; Mann, E.; Edwards, J.; Brown, D. T. Effects of chloroquine and cytochalasin B on the infection of cells by Sindbis virus and vesicular stomatitis virus. *J. Virol.* **1981**, 37, 1060-1065.
111. Helenius, A.; Marsh, M.; White, J. Inhibition of Semliki forest virus penetration by lysosomotropic weak bases. *J. Gen. Virol.* **1982**, 58, 47-61.
112. Cassell, S.; Edwards, J.; Brown, D. T. Effects of lysosomotropic weak bases on infection of BHK-21 cells by Sindbis virus. *J. Virol.* **1984**, 52, 857-864.
113. Maheshwari, R. K.; Srikantan, V.; Bhartiya, D. Chloroquine enhances replication of Semliki Forest virus and encephalomyocarditis virus in mice. *J. Virol.* **1991**, 65, 992-995.
114. Brighton, S. W. Chloroquine phosphate treatment of chronic Chikungunya arthritis. An open pilot study. *S. Afr. Med. J.* **1984**, 66, 217-218.
115. Leyssen, P.; De Clercq, E.; Neyts, J. The anti-yellow fever virus activity of ribavirin is independent of error-prone replication. *Mol. Pharmacol.* **2006**, 69, 1461-1467.
116. Briolant, S.; Garin, D.; Scaramozzino, N.; Jouan, A.; Crance, J. M. In vitro inhibition of chikungunya and Semliki Forest viruses replication by antiviral compounds: synergistic effect of interferon-alpha and ribavirin combination. *Antiviral Res.* **2004**, 61, 111-117.
117. Rada, B.; Dragun, M. Antiviral action and selectivity of 6-azauridine. *Ann. N. Y. Acad. Sci.* **1977**, 284, 410-417.
118. Crutcher, W. A.; Moschella, S. L. Double-blind controlled crossover high-dose study of Azaribine in psoriasis. *Br. J. Dermatol.* **1975**, 92, 199-205.
119. Panisheva, E. K. F., A. N.; Nikolaeva, I. S.; Galenko-Yaroshevskii, P. A.; Bartashevich, V. V.; Cherkasova, A. A.; Linchenko, S. N.; Egik'yan, A. L.;

Golovanova, E. A.; Pushkina, T. V. . Synthesis and biological activity of substituted 5-hydroxy-6-bromoindoles. *Khimiko-Farmatsevticheskii Zhurnal* **1988**, 27, 22-34.

120. Boriskin, Y. S.; Leneva, I. A.; Pecheur, E. I.; Polyak, S. J. Arbidol: A broad-spectrum antiviral compound that blocks viral fusion. *Curr. Med. Chem.* **2008**, 15, 997-1005.
121. Villalain, J. Membranotropic effects of arbidol, a broad anti-viral molecule, on phospholipid model membranes. *J. Phys. Chem. B* **2010**, 114, 8544-8554.
122. Leneva, I. A.; Russell, R. J.; Boriskin, Y. S.; Hay, A. J. Characteristics of arbidol-resistant mutants of influenza virus: Implications for the mechanism of anti-influenza action of arbidol. *Antiviral Res.* **2009**, 81, 132-140.
123. Pecheur, E. I.; Lavillette, D.; Alcaras, F.; Molle, J.; Boriskin, Y. S.; Roberts, M.; Cosset, F. L.; Polyak, S. J. Biochemical mechanism of hepatitis C virus inhibition by the broad-spectrum antiviral arbidol. *Biochemistry* **2007**, 46, 6050-6059.
124. Delogu, I.; Pastorino, B.; Baronti, C.; Nougairede, A.; Bonnet, E.; de Lamballerie, X. In vitro antiviral activity of arbidol against chikungunya virus and characteristics of a selected resistant mutant. *Antiviral Res.* **2011**, 90, 99-107.
125. Bentley, R. Mycophenolic Acid: a one hundred year odyssey from antibiotic to immunosuppressant. *Chem. Rev.* **2000**, 100, 3801-3826.
126. Chen, L.; Pankiewicz, K. W. Recent development of IMP dehydrogenase inhibitors for the treatment of cancer. *Curr. Opin. Drug Discov. Devel.* **2007**, 10, 403-412.
127. Ratcliffe, A. J. Inosine 5'-monophosphate dehydrogenase inhibitors for the treatment of autoimmune diseases. *Curr. Opin. Drug Discov. Devel.* **2006**, 9, 595-605.
128. Khan, M.; Dhanwani, R.; Patro, I. K.; Rao, P. V. L.; Parida, M. M. Cellular IMPDH enzyme activity is a potential target for the inhibition of chikungunya virus replication and virus induced apoptosis in cultured mammalian cells. *Antiviral Res.* **2011**, 89, 1-8.
129. Sweeney, M. J.; Hoffman, D. H.; Esterman, M. A. Metabolism and biochemistry of mycophenolic acid. *Cancer Res.* **1972**, 32, 1803-1809.
130. Bopp, R. J.; Schirmer, R. E.; Meyers, D. B. Determination of mycophenolic acid and its glucuronide metabolite in plasma. *J. Pharm. Sci.* **1972**, 61, 1750-1753.

131. Fan, D.; Ju, H.; Shao-Yong, Y.; Xia, C.; Da-Cai, Z.; Ye-Na, T. Trigonostemon tuberculatus (Euphorbiaceae), a peculiar new species from Yunnan Province, China. *Kew Bull.* **2010**, 65, 111-113.
132. Lin, B. D.; Han, M. L.; Ji, Y. C.; Chen, H. D.; Yang, S. P.; Zhang, S.; Geng, M. Y.; Yue, J. M. Trigoxyphins A-G: diterpenes from Trigonostemon xyphophylloides. *J. Nat. Prod.* **2010**, 73, 1301-1305.
133. Dong, S. H.; Zhang, C. R.; Xu, C. H.; Ding, J.; Yue, J. M. Daphnane-type diterpenoids from Trigonostemon howii. *J. Nat. Prod.* **2011**, 74, 1255-1261.
134. Zhang, L.; Luo, R. H.; Wang, F.; Jiang, M. Y.; Dong, Z. J.; Yang, L. M.; Zheng, Y. T.; Liu, J. K. Highly functionalized daphnane diterpenoids from Trigonostemon thyrsoideum. *Org. Lett.* **2010**, 12, 152-155.
135. Zhang, L.; Luo, R. H.; Wang, F.; Dong, Z. J.; Yang, L. M.; Zheng, Y. T.; Liu, J. K. Daphnane diterpenoids isolated from Trigonostemon thyrsoideum as HIV-1 antivirals. *Phytochemistry* **2010**, 71, 1879-1883.
136. Allard, P. M.; Martin, M. T.; Tran Huu Dau, M. E.; Leyssen, P.; Gueritte, F.; Litaudon, M. Trigocherrin A, the first natural chlorinated daphnane diterpene orthoester from Trigonostemon cherrieri. *Org. Lett.* **2012**, 14, 342-345.
137. Allard, P. M.; Leyssen, P.; Martin, M. T.; Bourjot, M.; Dumontet, V.; Eydoux, C.; Guillemot, J. C.; Canard, B.; Poullain, C.; Gueritte, F.; Litaudon, M. Antiviral chlorinated daphnane diterpenoid orthoesters from the bark and wood of Trigonostemon cherrieri. *Phytochemistry* **2012**, 84, 160-168.
138. Bourjot, M.; Delang, L.; Nguyen, V. H.; Neyts, J.; Gueritte, F.; Leyssen, P.; Litaudon, M. Prostratin and 12-O-Tetradecanoylphorbol 13-acetate are potent and selective inhibitors of chikungunya virus replication. *J. Nat. Prod.* **2012**.
139. Hezareh, M. Prostratin as a new therapeutic agent targeting HIV viral reservoirs. *Drug News Perspect.* **2005**, 18, 496-500.
140. Chowdhury, M. I. H.; Koyanagi, Y.; Kobayashi, S.; Hamamoto, Y.; Yoshiyama, H.; Yoshida, T.; Yamamoto, N. The Phorbol Ester TBA Strongly Inhibits HIV-1-Induced syncytia formation but enhances virus production - possible involvement of protein kinase-C pathway. *Virology* **1990**, 176, 126-132.
141. Voet, D. V., J. G. In *In Biochimie; de Boeck, Ed.*, Université: Bruxelles: 2005; p 712.

142. Szallasi, Z.; Krsmanovic, L.; Blumberg, P. M. Nonpromoting 12-deoxyphorbol 13-esters inhibit phorbol 12-myristate 13-acetate induced tumor promotion in CD-1 mouse skin. *Cancer Res.* **1993**, 53, 2507-2512.
143. Bourjot, M.; Leyssen, P.; Eydoux, C.; Guillemot, J. C.; Canard, B.; Rasoanaivo, P.; Gueritte, F.; Litaudon, M. Chemical constituents of Anacolosa pervilleana and their antiviral activities. *Fitoterapia* **2012**, 83, 1076-1080.
144. Takeda, S.; Yajima, N.; Kitazato, K.; Unemi, N. Antitumor activities of harringtonine and homoharringtonine, cephalotaxus alkaloids which are active principles from plant by intraperitoneal and oral administration. *J. Pharmacobiodyn.* **1982**, 5, 841-847.
145. Kaur, P.; Thiruchelvan, M.; Lee, R. C.; Chen, H.; Chen, K. C.; Ng, M. L.; Chu, J. J. Inhibition of chikungunya virus replication by harringtonine, a novel antiviral that suppresses viral protein expression. *Antimicrob. Agents Chemother.* **2013**, 57, 155-167.
146. D'Hooghe, M.; Mollet, K.; De Vreese, R.; Jonckers, T. H.; Dams, G.; De Kimpe, N. Design, synthesis, and antiviral evaluation of purine-beta-lactam and purine-aminopropanol hybrids. *J. Med. Chem.* **2012**, 55, 5637-5641.
147. Field, A. K.; Tytell, A. A.; Lampson, G. P.; Hilleman, M. R. Inducers of interferon and host resistance. II. Multistranded synthetic polynucleotide complexes. *Proc. Natl. Acad. Sci. USA* **1967**, 58, 1004-1010.
148. Djeu, J. Y.; Heinbaugh, J. A.; Holden, H. T.; Herberman, R. B. Role of macrophages in the augementation of mouse natural killer cell activity by poly I:C and interferon. *J. Immunol.* **1979**, 122, 182-188.
149. Guillot, L.; Le Goffic, R.; Bloch, S.; Escriou, N.; Akira, S.; Chignard, M.; Si-Tahar, M. Involvement of toll-like receptor 3 in the immune response of lung epithelial cells to double-stranded RNA and influenza A virus. *J. Biol. Chem.* **2005**, 280, 5571-5580.
150. Li, Y. G.; Siripanyaphinyo, U.; Tumkosit, U.; Noranate, N.; A, A. N.; Pan, Y.; Kameoka, M.; Kurosu, T.; Ikuta, K.; Takeda, N.; Anantapreecha, S. Poly (I:C), an agonist of toll-like receptor-3, inhibits replication of the Chikungunya virus in BEAS-2B cells. *Virol. J.* **2012**, 9, 114.

151. Dash, P. K.; Tiwari, M.; Santhosh, S. R.; Parida, M.; Rao, P. V. L. RNA interference mediated inhibition of chikungunya virus replication in mammalian cells. *Biochem. Biophys. Res. Commun.* **2008**, 376, 718-722.
152. Lam, S.; Chen, K. C.; Ng, M. M.; Chu, J. J. Expression of plasmid-based shRNA against the E1 and nsP1 genes effectively silenced chikungunya virus replication. *PLoS One* **2012**, 7, e46396.
153. Luu, T. T. T. Study of reverse transcriptase: synthesis of inhibitors, conformational analysis, and integration of drug design. PhD dissertation. University of Wollongong, Wollongong, 2000.
154. Keller, P. A.; Birch, C.; Leach, S. P.; Tyssen, D.; Griffith, R. Novel pharmacophore-based methods reveal gossypol as a reverse transcriptase inhibitor. *J. Mol. Graphics Model.* **2003**, 21, 365-373.
155. Yepuri, N. R.; Haritakul, R.; Griffith, R.; Leach, S. P.; Keller, P. A. The synthesis and testing of arenearylpyrimidylmethanes as antimalarial agents. *Chemmedchem* **2006**, 1, 715-717.
156. Yepuri, N. R. The design and synthesis of novel anti-malarial agents. PhD dissertation. University of Wollongong, Wollongong, 2004.
157. Stevens, A. J. The Design, synthesis and evaluation of potential inhibitors targeting malaria and dengue fever virus. PhD Dissertation. University of Wollongong, Wollongong, 2011.
158. Rigau-Perez, J. G.; Clark, G. G.; Gubler, D. J.; Reiter, P.; Sanders, E. J.; Vorndam, A. V. Dengue and dengue haemorrhagic fever. *Lancet* **1998**, 352, 971-977.
159. Cole, J. G.; Wan, P. Mechanistic studies of photohydration of m-hydroxy-1,1-diaryl alkenes. *Can. J. Chem.* **2002**, 80, 46-54.
160. Campaign. E; Ho, J.; Bradford, M. Synthesis of Some Amino-substituted furanopyrimidines and styrylpyrimidines. *J. Heterocyclic Chem.* **1970**, 7, 257-260.
161. Jalilzadeh, M.; Pesyan, N. N. New strategy for the synthesis of 5-aryl-1H,1' H-spiro[furo[2,3-*d*]pyrimidine-6,5'-pyrimidine]2,2 ',4,4',6' (3H, 3'H, 5H)-pentaones and their sulfur analogues. *Bull. Korean Chem. Soc.* **2011**, 32, 3382-3388.
162. Polshettiwar, V.; Varma, R. S. Microwave-assisted organic synthesis and transformations using benign reaction media. *Acc. Chem. Res.* **2008**, 41, 629-639.

163. Radinov, R.; Haimova, M.; Simova, E. Synthesis of 4-amino-3-pyridinyl and 4-amino-5-pyrimidinyl aryl ketones and related-compounds via an ortho-lithiation reaction. *Synthesis* **1986**, 886-891.
164. Vicha, R.; Kuritka, I.; Rouchal, M.; Jezkova, V.; Zierhut, A. Directing effects in nitration of 1-adamantyl bearing aromatic ketones. *Arkivoc* **2009**, 60-80.
165. Margolis, B. J.; Long, K. A.; Laird, D. L.; Ruble, J. C.; Pulley, S. R. Assembly of 4-aminoquinolines via palladium catalysis: a mild and convenient alternative to S_NAr methodology. *J. Org. Chem.* **2007**, 72, 2232-2235.
166. Csuk, R.; Barthel, A.; Raschke, C.; Kluge, R.; Strohl, D.; Trieschmann, L.; Bohm, G. Synthesis of monomeric and dimeric acridine compounds as potential therapeutics in alzheimer and prion diseases. *Arch. Pharm.* **2009**, 342, 699-709.
167. Benaskar, F.; Engels, V.; Patil, N.; Rebrov, E. V.; Meuldijk, J.; Hessel, V.; Hulshof, L. A.; Jefferson, D. A.; Schouten, J. C.; Wheatley, A. E. H. Copper(0) in the Ullmann heterocycle-aryl ether synthesis of 4-phenoxypyridine using multimode microwave heating (vol 51, pg 248, 2010). *Tetrahedron Lett.* **2010**, 51, 5849-5849.
168. Barta, T. E. B., Daniel P.; Bedell, Louis J.; Boehm, Terri L.; Carroll, Jeffrey N.; Decrescenzo, Gary A.; Fobian, Yvette M.; Freskos, John N.; Getman, Daniel P.; McDonald, Joseph J.; Hockerman, Susan L.; Howard, Susan C.; Kolodziej, Stephen A.; Li, Madeleine Hui; Mischke, Deborah A.; Rico, Joseph G.; Stehle, Nathan W.; Tollefson, Michael B.; Vernier, William F.; Villamil, Clara I. . Aromatic sulfone hydroxamic acid metalloprotease inhibitor. US 6750228 B1, **2004**.
169. Kumar, M.; Sharma, U.; Sharma, S.; Kumar, V.; Singh, B.; Kumar, N. Catalyst-free water mediated reduction of nitroarenes using glucose as a hydrogen source. *RSC Adv.* **2013**, 3, 4894-4898.
170. Gkizis, P. L.; Stratakis, M.; Lykakis, I. N. Catalytic activation of hydrazine hydrate by gold nanoparticles: Chemoselective reduction of nitro compounds into amines. *Catal. Commun.* **2013**, 36, 48-51.
171. Shi, Q. X.; Lu, R. W.; Zhang, Z. X.; Zhao, D. F. Preparation of sulphur-containing aromatic amines by reduction of the corresponding aromatic nitro compounds with hydrazine hydrate over iron oxide hydroxide catalyst. *Chinese Chem. Lett.* **2006**, 17, 1045-1047.

172. Trott, O.; Olson, A. J. Autodock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, 31, 455-461.
173. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, 30, 2785-2791.
174. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, 19, 1639-1662.
175. Durrant, J. D.; Urbaniak, M. D.; Ferguson, M. A. J.; McCammon, J. A. Computer-aided identification of trypanosoma brucei uridine diphosphate galactose 4 '-epimerase inhibitors: toward the development of novel therapies for african sleeping sickness. *J. Med. Chem.* **2010**, 53, 5025-5032.
176. Schwobel, J.; Ebert, R. U.; Kuhne, R.; Schuurmann, G. Prediction of the intrinsic hydrogen bond acceptor strength of organic compounds by local molecular parameters. *J. Chem. Inf. Model.* **2009**, 49, 956-962.
177. Lansdon, E. B.; Brendza, K. M.; Hung, M.; Wang, R.; Mukund, S.; Jin, D.; Birkus, G.; Kutty, N.; Liu, X. Crystal structures of HIV-1 reverse transcriptase with etravirine (TMC125) and rilpivirine (TMC278): implications for drug design. *J. Med. Chem.* **2010**, 53, 4295-4299.
178. Ponnala, S.; Prasad Sahu, D. Iodine-mediated synthesis of 2-arylbenzoxazoles, 2-arylbenzimidazoles, and 1,3,5-trisubstituted pyrazoles. *Synthetic Commun.* **2006**, 36, 2189-2194.
179. Wang, R.; Lu, X.-x.; Yu, X.-q.; Shi, L.; Sun, Y. Acid-catalyzed solvent-free synthesis of 2-arylbenzimidazoles under microwave irradiation. *J. Mol. Catal. A: Chemical* **2007**, 266, 198-201.
180. Heravi, M. M.; Tajbakhsh, M.; Ahmadi, A. N.; Mohajerani, B. Zeolites. Efficient and eco-friendly catalysts for the synthesis of benzimidazoles. *Monatsh. Chem.* **2006**, 137, 175-179.
181. Lee, J. P.; Bembi, R.; Fife, T. H. Steric Effects in the Hydrolysis Reactions of N-Acylimidazoles. Effect of aryl substitution in the leaving group. *J. Org. Chem.* **1997**, 62, 2872-2876.

182. Fife, T. H.; Natarajan, R.; Werner, M. H. Effect of the leaving group in the hydrolysis of *N*-acylimidazoles. The hydroxide ion, water, and general-base catalyzed hydrolysis of *N*-acyl-4(5)-nitroimidazoles. *J. Org. Chem.* **1987**, 52, 740-746.
183. Ando, K.; Suzuki, S.; Arita, M. Synthesis of mycophenolic acid beta-D-glucuronide and its antitumor activity. *J. Antibiot.* **1970**, 23, 408-413.
184. Nery, R.; Nice, E. The metabolism and binding of (14C)mycophenolic acid in the rat. *J. Pharm. Pharmacol.* **1971**, 23, 842-847.
185. Bhal, S. K.; Kassam, K.; Peirson, I. G.; Pearl, G. M. The rule of five revisited: applying LogD in place of logP in drug-likeness filters. *Mol. Pharm.* **2007**, 4, 556-560.
186. Tetko, I. V.; Poda, G. I. Application of ALOGPS 2.1 to predict log D distribution coefficient for Pfizer proprietary compounds. *J. Med. Chem.* **2004**, 47, 5601-5604.
187. da Silva, J. F. M.; Garden, S. J.; Pinto, A. C. The chemistry of isatins: a review from 1975 to 1999. *J. Brazil. Chem. Soc.* **2001**, 12, 273-286.
188. Candido-Bacani Pde, M.; Mori, M. P.; Calvo, T. R.; Vilegas, W.; Varanda, E. A.; de Syllos Colus, I. M. In vitro assessment of the cytotoxic, apoptotic, and mutagenic potentials of isatin. *J. Toxicol. Environ. Health A* **2013**, 76, 354-362.
189. Dischia, M.; Palumbo, A.; Prota, G. adrenaline oxidation revisited - new products beyond the adrenochrome stage. *Tetrahedron* **1988**, 44, 6441-6446.
190. Medvedev, A. E.; Clow, A.; Sandler, M.; Glover, V. Isatin: a link between natriuretic peptides and monoamines? *Biochem. Pharmacol.* **1996**, 52, 385-391.
191. Binda, C.; Li, M.; Hubalek, F.; Restelli, N.; Edmondson, D. E.; Mattevi, A. Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 9750-9755.
192. Cao, J. R.; Gao, H.; Bemis, G.; Salituro, F.; Ledeboer, M.; Harrington, E.; Wilke, S.; Taslimi, P.; Pazhanisamy, S.; Xie, X. L.; Jacobs, M.; Green, J. Structure-based design and parallel synthesis of N-benzyl isatin oximes as JNK3 MAP kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, 19, 2891-2895.
193. Hubalek, F.; Binda, C.; Khalil, A.; Li, M.; Mattevi, A.; Castagnoli, N.; Edmondson, D. E. Demonstration of isoleucine 199 as a structural determinant for

the selective inhibition of human monoamine oxidase B by specific reversible inhibitors. *J. Biol. Chem.* **2005**, 280, 15761-15766.

194. Miyaura, N. Heck and cross-coupling reactions: Two core chemistries in metal-catalyzed organic syntheses. *Adv. Synth. Catal.* **2004**, 346, 1522-1523.
195. Yin, L. X.; Liebscher, J. Carbon-carbon coupling reactions catalyzed by heterogeneous palladium catalysts. *Chem. Rev.* **2007**, 107, 133-173.
196. Alonso, F.; Beletskaya, I. P.; Yus, M. Non-conventional methodologies for transition-metal catalysed carbon-carbon coupling: a critical overview. Part 2: The Suzuki reaction. *Tetrahedron* **2008**, 64, 3047-3101.
197. Biot, C.; Bauer, H.; Schirmer, R. H.; Davioud-Charvet, E. 5-Substituted tetrazoles as bioisosteres of carboxylic acids. Bioisosterism and mechanistic studies on glutathione reductase inhibitors as antimalarials. *J. Med. Chem.* **2004**, 47, 5972-5983.
198. Patrick, G. L. *An Introduction to Medicinal Chemistry* 4th ed.; Oxford University Press: 2009.
199. Liljebris, C.; Larsen, S. D.; Ogg, D.; Palazuk, B. J.; Bleasdale, J. E. Investigation of potential bioisosteric replacements for the carboxyl groups of peptidomimetic inhibitors of protein tyrosine phosphatase 1B: Identification of a tetrazole-containing inhibitor with cellular activity. *J. Med. Chem.* **2002**, 45, 1785-1798.
200. Garcia-Munoz, A. H.; Tomas-Gamasa, M.; Perez-Aguilar, M. C.; Cuevas-Yanez, E.; Valdes, C. Straightforward reductive esterification of carbonyl compounds with carboxylic acids through tosylhydrazone intermediates. *Eur. J. Org. Chem.* **2012**, 3925-3928.
201. Shirakawa, S.; Takai, J.; Sasaki, K.; Miura, T.; Maruoka, K. Bowl-shaped [tris(2,6-diphenylbenzyl)siloxy]dimethylaluminum catalyst for effecting Tishchenko reaction. *Heterocycles* **2003**, 59, 57-61.
202. Endo, Y.; Backvall, J. E. Aerobic Lactonization of diols by biomimetic Oxidation. *Chem-Eur. J.* **2011**, 17, 12596-12601.
203. Zhang, Y. H.; Shi, B. F.; Yu, J. Q. Palladium(II)-catalyzed ortho alkylation of benzoic acids with alkyl halides. *Angew. Chem. Int. Edit.* **2009**, 48, 6097-6100.
204. Allison W. R., N. G. T. Lactones. Part VI. The Preparation of 5,7-dihydroxyphthalide, its 5- methyl ether, and related compounds. *J. Chem. Soc.* **1979**, 3335-3340.

205. Jongen, R.; Sala, T.; Sargent, M. V. Depsidone synthesis . Part 13. Total synthesis of variolaric acid. *J. Chem. Soc. Perk. Trans. 1* **1979**, 2588-2592.
206. Mirabdolbaghi, R.; Dudding, T. An Indium-mediated allylative/ transesterification DFT-directed approach to chiral C-(3)-functionalized phthalides. *Org. Lett.* **2012**, 14, 3748-3751.
207. Lee, Y. M.; Fujiwara, Y.; Ujita, K.; Nagatomo, M.; Ohta, H.; Shimizu, I. Syntheses of mycophenolic acid and its analogs by palladium methodology. *Bull. Chem. Soc. Jpn.* **2001**, 74, 1437-1443.
208. Wang, J. Y.; Johnson, D. M. Design, synthesis and polymerization of highly branched pseudodendrimers through tandem reactions. *Polym. Int.* **2009**, 58, 1234-1245.
209. Gamble, A. B.; Garner, J.; Gordon, C. P.; O'Conner, S. M. J.; Keller, P. A. Aryl nitro reduction with iron powder or stannous chloride under ultrasonic irradiation. *Synthetic Commun.* **2007**, 37, 2777-2786.
210. Pokhodylo, N. T.; Matiychuk, V. S.; Obushak, M. D. Synthesis of isothiocoumarin derivatives. *Chem. Heterocycl. Com.* **2010**, 46, 140-145.
211. Loloiu, G.; Maior, O. Isatin chemistry. Synthesis of *N*-methyl-2,3-dioxo-2,3-dihydropyrrolo[2,3-*b*]phenoxatiin. *Rev. Roum. Chim.* **1997**, 42, 67-69.
212. Gungor, T.; Chen, Y.; Golla, R.; Ma, Z. P.; Corte, J. R.; Northrop, J. P.; Bin, B.; Dickson, J. K.; Stouch, T.; Zhou, R.; Johnson, S. E.; Seethala, R.; Feyen, J. H. M. Synthesis and characterization of 3-arylquinazolinone and 3-arylquinazolinethione derivatives as selective estrogen receptor beta modulators. *J. Med. Chem.* **2006**, 49, 2440-2455.
213. Gopal, M.; Srivastava, G.; Pande, U. C.; Tiwari, R. D. Microdetermination of indoles with *N*-Bromoacetamide in acetic-acid medium. *Mikrochim. Acta* **1977**, 2, 215-220.
214. Sandmeyer, T. Über Isonitrosoacetanilide und deren Kondensation zu Isatinen. *Helv. Chim. Acta* **1919**, 2, 234–242.
215. Montoya-Pelaez, P. J.; Uh, Y. S.; Lata, C.; Thompson, M. P.; Lemieux, R. P.; Crudden, C. M. The synthesis and resolution of 2,2'-, 4,4'-, and 6,6'-substituted chiral biphenyl derivatives for application in the preparation of chiral materials. *J. Org. Chem.* **2006**, 71, 5921-5929.

216. Fischer, A.; Henderson, G. N.; Mahasay, S. R. Ipso nitration . XXIX. Nitration of Substituted 4-Methylanisoles and Phenols. *Can. J. Chem.* **1987**, 65, 1233-1240.
217. Ragan, J. A.; Makowski, T. W.; Castaldi, M. J.; Hill, P. D. 2,5-Dimethylpyrrole protection facilitates copper(I)-mediated methoxylations of aryl iodides in the presence of anilines. *Synthesis* **1998**, 1599-1603.
218. Ishiyama, T.; Murata, M.; Miyaura, N. Palladium(0)catalyzed cross-coupling reaction of alkoxydiboron with haloarenes - a direct procedure for arylboronic esters. *J. Org. Chem.* **1995**, 60, 7508-7510.
219. Himo, F.; Demko, Z. P.; Noodleman, L.; Sharpless, K. B. Mechanisms of tetrazole formation by addition of azide to nitriles. *J. Am. Chem. Soc.* **2002**, 124, 12210-12216.
220. Caron, S.; Wei, L. L.; Douville, J.; Ghosh, A. Preparation and utility of trihaloethyl imidates: useful reagents for the synthesis of amidines. *J. Org. Chem.* **2010**, 75, 945-947.
221. Stokes, B. J.; Vogel, C. V.; Urnezis, L. K.; Pan, M.; Driver, T. G. Intramolecular Fe(II)-catalyzed N-O or N-N bond formation from aryl azides. *Org. Lett.* **2010**, 12, 2884-2887.
222. Dai, W. M.; Cheung, Y. K.; Tang, K. W.; Choi, P. Y.; Chung, S. L. Highly chemoselective acylation of substituted aminophenols with 3-(trimethylacetyl)-1,3-thiazolidine-2-thione. *Tetrahedron* **1995**, 51, 12263-12276.
223. Blunk, D.; Porada, J. H. Liquid crystallinity and supramolecular organization of regioisomeric isatin derivatives: an experimental and theoretical study. *Chemphyschem* **2009**, 10, 3260-3264.
224. Porada, J. H.; Neudorfl, J.; Blunk, D. Synthesis and supramolecular organization of 5-(4-alkylphenyl)isatin. *Cryst. Growth. Des.* **2011**, 11, 3648-3652.
225. Tamm, K.; Merits, A.; Sarand, I. Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. *J. Gen. Virol.* **2008**, 89, 676-686.
226. McInnes, C. Virtual screening strategies in drug discovery. *Curr. Opin. Chem. Biol.* **2007**, 11, 494-502.
227. Kirchmair, J.; Distinto, S.; Schuster, D.; Spitzer, G.; Langer, T.; Wolber, G. Enhancing drug discovery through in silico screening: Strategies to increase true positives retrieval rates. *Curr. Med. Chem.* **2008**, 15, 2040-2053.

228. Yennamalli, R.; Subbarao, N.; Kampmann, T.; McGeary, R. P.; Young, P. R.; Kobe, B. Identification of novel target sites and an inhibitor of the dengue virus E protein. *J. Comput.-Aided Mol. Des.* **2009**, 23, 333-341.
229. Zhou, Z. G.; Khaliq, M.; Suk, J. E.; Patkar, C.; Li, L.; Kuhn, R. J.; Post, C. B. Antiviral compounds discovered by virtual screening of small-molecule libraries against dengue virus E protein. *ACS Chem. Biol.* **2008**, 3, 765-775.
230. Li, Z.; Khaliq, M.; Zhou, Z. G.; Post, C. B.; Kuhn, R. J.; Cushman, M. Design, synthesis, and biological evaluation of antiviral agents targeting flavivirus envelope proteins. *J. Med. Chem.* **2008**, 51, 4660-4671.
231. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Del. Rev.* **2001**, 46, 3-26.
232. McGann, M. FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, 51, 578-596.
233. McGaughey, G. B.; Sheridan, R. P.; Bayly, C. I.; Culberson, J. C.; Kreatsoulas, C.; Lindsley, S.; Maiorov, V.; Truchon, J. F.; Cornell, W. D. Comparison of topological, shape, and docking methods in virtual screening. *J. Chem. Inf. Model.* **2007**, 47, 1504-1519.
234. McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian docking functions. *Biopolymers* **2003**, 68, 76-90.
235. Bostrom, J.; Greenwood, J. R.; Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graphics Model.* **2003**, 21, 449-462.
236. Hawkins, P. C.; Nicholls, A. Conformer generation with OMEGA: learning from the data set and the analysis of failures. *J. Chem. Inf. Model.* **2012**, 52, 2919-2936.
237. Mayhoub, A. S.; Khaliq, M.; Kuhn, R. J.; Cushman, M. Design, synthesis, and biological evaluation of thiazoles targeting flavivirus envelope proteins. *J. Med. Chem.* **2011**, 54, 1704-1714.
238. Cecchi, G.; Paone, M.; Franco, J. R.; Fevre, E. M.; Diarra, A.; Ruiz, J. A.; Mattioli, R. C.; Simarro, P. P. Towards the Atlas of human african trypanosomiasis. *Int. J. Health. Geogr.* **2009**, 8, 10.1186/1476-072X-8-15.
239. Brun, R.; Blum, J.; Chappuis, F.; Burri, C. Human african trypanosomiasis. *Lancet* **2010**, 375, 148-159.

240. Popejoy, M. W. Working to overcome the global impact of neglected tropical diseases. *Perspect. Public Heal.* **2012**, 132, 192-192.
241. Wastling, S. L.; Welburn, S. C. Diagnosis of human sleeping sickness: sense and sensitivity. *Trends Parasitol.* **2011**, 27, 394-402.
242. Phillips, M. A. Stoking the drug target pipeline for human African trypanosomiasis. *Mol. Microbiol.* **2012**.
243. Michael, C.; Turner, R. Antigenic variation in trypanosoma brucei infections: an holistic view. *J. Cell Sci.* **1999**, 112, 3187-3192.
244. Croft, S. L.; Barrett, M. P.; Urbina, J. A. Chemotherapy of trypanosomiases and leishmaniasis. *Trends Parasitol.* **2005**, 21, 508-512.
245. Renslo, A. R.; McKerrow, J. H. Drug discovery and development for neglected parasitic diseases. *Nat. Chem. Biol.* **2006**, 2, 701-710.
246. Pepin, J.; Milord, F. The Treatment of human african trypanosomiasis. *Adv. Parasitol.* **1994**, 33, 1-47.
247. Barrett, M. P. Potential new drugs for human african trypanosomiasis: some progress at last. *Curr. Opin. Infect. Dis.* **2010**, 23, 603-608.
248. Priotto, G.; Kasparian, S.; Mutombo, W.; Ngouama, D.; Ghorashian, S.; Arnold, U.; Ghabri, S.; Baudin, E.; Buard, V.; Kazadi-Kyanza, S.; Ilunga, M.; Mutangala, W.; Pohlig, G.; Schmid, C.; Karunakara, U.; Torreele, E.; Kande, V. Nifurtimox-eflornithine combination therapy for second-stage african trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet* **2009**, 374, 56-64.
249. Yun, O.; Priotto, G.; Tong, J.; Flevaud, L.; Chappuis, F. NECT is next: Implementing the new drug combination therapy for trypanosoma brucei gambiense sleeping sickness. *Plos Neglect. Trop. Dis.* **2010**, 4(5), e720.
250. Delespaux, V.; de Koning, H. P. Drugs and drug resistance in african trypanosomiasis. *Drug Resist. Update* **2007**, 10, 30-50.
251. Rodgers, J. Human African trypanosomiasis, chemotherapy and CNS disease. *J. Neuroimmunol.* **2009**, 211, 16-22.
252. Wenzler, T.; Boykin, D. W.; Ismail, M. A.; Hall, J. E.; Tidwell, R. R.; Brun, R. New treatment option for second-stage african sleeping sickness: In vitro and in vivo efficacy of aza analogs of DB289. *Antimicrob. Agents Chemother.* **2009**, 53, 4185-4192.

253. Torreele, E.; Trunz, B. B.; Tweats, D.; Kaiser, M.; Brun, R.; Mazue, G.; Bray, M. A.; Pecoul, B. Fexinidazole - a new oral nitroimidazole drug candidate entering clinical development for the treatment of sleeping sickness. *Plos Neglect. Trop. D* **2010**, 4(12), e923.

254. Jacobs, R. T.; Nare, B.; Wring, S. A.; Orr, M. D.; Chen, D.; Sligar, J. M.; Jenks, M. X.; Noe, R. A.; Bowling, T. S.; Mercer, L. T.; Rewerts, C.; Gaukel, E.; Owens, J.; Parham, R.; Randolph, R.; Beaudet, B.; Bacchi, C. J.; Yarlett, N.; Plattner, J. J.; Freund, Y.; Ding, C.; Akama, T.; Zhang, Y. K.; Brun, R.; Kaiser, M.; Scandale, I.; Don, R. SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human african trypanosomiasis. *Plos Neglect. Trop. D* **2011**, 5(6), e1151.

255. Ettari, R.; Tamborini, L.; Angelo, I. C.; Micale, N.; Pinto, A.; De Micheli, C.; Conti, P. Inhibition of rhodesain as a novel therapeutic modality for human african trypanosomiasis. *J. Med. Chem.* **2013**, 56, 5637-5658

256. Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, 53, 2719-2740.

257. Uchiyama, M.; Matsumoto, Y.; Nakamura, S.; Ohwada, T.; Kobayashi, N.; Yamashita, N.; Matsumiya, A.; Sakamoto, T. Development of a catalytic electron transfer system mediated by transition metal ate complexes: applicability and tunability of electron-releasing potential for organic transformations. *J. Am. Chem. Soc.* **2004**, 126, 8755-8759.

258. Pagliero, R. J.; Pierini, A. B.; Brun, R.; Mazzieri, M. R. Design, synthesis and 3-D characterization of 1-benzenesulfonyl-1,2,3,4-tetrahydroquinolinesas lead scaffold for antiparasitic drug. *Lett. Drug. Des. Discov.* **2010**, 7, 461-470.

259. Buchstaller, H. P.; Siebert, C. D.; Steinmetz, R.; Frank, I.; Berger, M. L.; Gottschlich, R.; Leibrock, J.; Krug, M.; Steinhilber, D.; Noe, C. R. Synthesis of thieno[2,3-*b*]pyridinones acting as cytoprotectants and as inhibitors of [H-3] Glycine binding to the *N*-methyl-D-aspartate (NMDA) receptor. *J. Med. Chem.* **2006**, 49, 864-871.

260. Hwang, K. J.; Lee, T. S.; Kim, K. W.; Kim, B. T.; Lee, C. M.; Park, E. Y.; Woo, R. S. 4-hydroxy-6-oxo-6,7-dihydro-thieno[2,3-*b*] pyrimidine derivatives: Synthesis and their biological evaluation for the glycine site acting on the *N*-methyl-D-aspartate (NMDA) receptor. *Arch. Pharm. Res.* **2001**, 24, 270-275.

261. Martinez-Viturro, C. M.; Dominguez, D. Synthesis of aza analogues of the anticancer agent batracylin. *Tetrahedron Lett.* **2007**, 48, 4707-4710.

262. Rashad, A. A.; Keller, P. A. Structure based design towards the identification of novel binding sites and inhibitors for the chikungunya virus envelope proteins. *J. Mol. Graphics Model.* **2013**, 44, 241-252.

263. Bennett, J. G.; Bunce, S. C. Cyclopropyl analogs of hexestrol and diethylstilbestrol. *J. Org. Chem.* **1960**, 25, 73-79.

264. Mali, R. S.; Jagtap, P. G.; Tilve, S. G. Convenient synthesis of naturally-occurring methoxy phthalides and hydroxy phthalides. *Synthetic Commun.* **1990**, 20, 2641-2652.

265. Christiansen, E.; Due-Hansen, M. E.; Urban, C.; Grundmann, M.; Schmidt, J.; Hansen, S. V.; Hudson, B. D.; Zaibi, M.; Markussen, S. B.; Hagesaether, E.; Milligan, G.; Cawthorne, M. A.; Kostenis, E.; Kassack, M. U.; Ulven, T. Discovery of a potent and selective free fatty acid receptor 1 agonist with low lipophilicity and high oral bioavailability. *J. Med. Chem.* **2013**, 56, 982-992.

266. Marshall Kulka, R. H. F. M. The nitration of some quinoline derivatives. *Can. J. Chem.* **1952**, 30, 720-724.

267. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. Uff, a full periodic-table force-field for molecular mechanics and molecular-dynamics simulations. *J. Am. Chem. Soc.* **1992**, 114, 10024-10035.